Direct Assimilation of All-Sky GOES-R ABI Radiances in GSI EnKF for the Analysis and Forecasting of a Mesoscale Convective System

LIJIAN ZHU,[a,b,d] MING XUE,[a,b,c] RONG KONG,[b] AND JINZHONG MIN[a]

[a] *Collaborative Innovation Center on Forecast and Evaluation of Meteorological Disasters, Key Laboratory of Meteorological Disaster of Education, Nanjing University of Information Science and Technology, Nanjing, China*
[b] *Center for Analysis and Prediction of Storms, University of Oklahoma, Norman, Oklahoma*
[c] *School of Meteorology, University of Oklahoma, Norman, Oklahoma*
[d] *Shanghai Typhoon Institute of China Meteorological Administration, Shanghai, China*

(Manuscript received 9 November 2021, in final form 1 October 2022)

ABSTRACT: In this study, all-sky GOES-R ABI infrared radiances at their native resolution are assimilated using an enhanced GSI ensemble Kalman filter (EnKF) data assimilation (DA) system, and the impacts of the data on the analysis and forecast of a mesoscale convective system (MCS) are explored. Results show that all-sky ABI BT data can correctly build up observed storms within the model and effectively remove spurious storms in model background through frequent DA cycles. Both bias and root-mean-squared innovation of the background and analysis are significantly reduced during the DA cycles, and free forecasts are improved when verified subjectively and objectively against observed ABI BTs and independent radar reflectivity observations. A horizontal localization radius of 30 km is found to produce the best results while 5-min DA cycles improve the storm analyses over 15-min cycles, but the differences in forecasts are small. Further analyses show that the clearing of spurious clouds by ABI radiance is correctly accompanied by reduction in moisture through background error cross covariance, but overdrying often occurs, which can cause spurious storm decay in the forecast. The problem is reduced when the ensemble mean of observation prior instead of observation prior of the ensemble mean state is used in the ensemble mean state update equation of EnKF. The significant difference between the two ways that the ensemble mean of observation prior is calculated when the observational operator is very nonlinear has not been recognized in earlier cloudy radiance DA studies.

SIGNIFICANCE STATEMENT: Satellite observations in cloudy regions are not used in most current operational weather prediction systems due to complex nonlinear relations between satellite-observed quantities, the radiances, and model state in such regions. The models also must predict clouds reasonably well for cloudy observations to be effectively assimilated. The latest NOAA geostationary satellites can provide radiance observations at high spatial and temporal resolutions and such data in both cloudy- and clear-air regions are assimilated using an advanced data assimilation method into a model that explicitly represents convection. Forecasts up to 4 h are improved by the assimilation while several issues associated with the assimilation are discussed. The study contributes to the eventual use of all-sky satellite radiance data in operational models.

KEYWORDS: Satellite observations; Numerical weather prediction/forecasting; Cloud resolving models; Data assimilation; Ensembles

1. Introduction

High-impact weather events, such as hurricanes, severe thunderstorms, and tornadoes, cause significant property loss and fatalities. The short-range prediction of these weather events is, therefore, important but still very challenging mostly because of their rapid evolution and insufficient accuracy in the initial conditions of numerical weather prediction (NWP) models. Where Doppler weather radar data are available, the assimilation of radar reflectivity and radial velocity observations provides important information on convective storms and precipitation systems, and has been shown to improve short-range forecasting of high-impact weather (e.g., Kain et al. 2010; Stensrud et al. 2013; Sun et al. 2014). However, operational Doppler weather radars typically do not measure non-precipitating clouds or the environment outside storms. Also, the coverage of operational weather radars is usually limited to regions over land and even over land there are often measurement gaps due to blockages, etc. In comparison, satellite observations usually have much wider spatial coverage and can provide information on atmospheric environmental temperature, moisture and cloud properties prior to the formation of precipitation particles that can be detected by weather radars.

With the aid of radiative transfer models (RTMs), a large number of infrared (IR) and microwave (MW) satellite radiance data have already been directly assimilated into global as well as regional models, and are proven to yield great benefits (Eyre et al. 1993; Derber and Wu 1998; McNally et al. 2006; Pavelin et al. 2008; McNally 2009; Geer et al. 2010; Okamoto 2013; Zhu et al. 2016; Geer et al. 2018). Compared to geostationary satellites, data from polar-orbiting satellites have shown larger impacts in global models. However, instruments on board

Corresponding author: Ming Xue, mxue@ou.edu

DOI: 10.1175/MWR-D-21-0293.1

polar-orbiting satellites usually have low temporal resolutions (typically twice a day) that are insufficient for mesoscale or storm-scale applications. Geostationary instruments are able to observe Earth every few minutes with higher spatial resolutions (often less than 10 km); therefore, making it possible to provide nearly continuous coverage of the evolution of weather phenomena even at the convective scale (Stengel et al. 2009; Zou et al. 2011).

In recent years, the new-generation geostationary weather satellites, including *Himawari-8/9* operated by the Japan Meteorological Agency (Bessho et al. 2016), the GOES-R series operated by the U.S. National Oceanic and Atmospheric Administration (Schmit et al. 2005) and the Fengyun-4 series operated by the China Meteorological Administration (J. Yang et al. 2017), were launched successively, carrying more advanced and upgraded instruments than their predecessors, including the infrared imagers the Advanced *Himawari-8/9* Imager (AHI), the Advanced Baseline Imager (ABI), and the Advanced Geosynchronous Radiation Imager (AGRI) on these respective satellites. Due to their much higher spatial and temporal resolutions, the assimilation of such observations has great potential to improve severe weather forecasting.

Observations and their derived products from geostationary satellites have been assimilated into NWP systems for several decades. Many studies have demonstrated the positive impacts of assimilating clear-sky geostationary IR radiances (e.g., Köpken et al. 2004; Szyndel et al. 2005; Montmerle et al. 2007; Zou et al. 2011; Qin et al. 2013; Zou et al. 2015; Ma et al. 2017; Qin et al. 2017; C. Yang et al. 2017; Wang et al. 2018) and more recently of satellite-derived cloud water path (CWP) products (Jones et al. 2013a; Chen et al. 2015; Jones and Stensrud 2015). However, direct assimilation of geostationary IR radiance data in cloudy regions and within convective scale models is still very limited. The short predictability of cloud features (Fabry and Sun 2010) and the strong nonlinear dependency of cloudy radiance on thermodynamic and hydrometeor profiles are some of the reasons. To effectively assimilate cloudy radiance, the forecast model that provides background has to be able to produce reasonable prediction of simulated radiance in cloudy regions.

Some progress has been made in the direct assimilation of cloudy radiance data very recently. At the convective scale, precipitation microphysics plays an important role, and cloud and hydrometeors are directly involved in the cloudy radiance calculation (e.g., Cintineo et al. 2014). Unlike variational data assimilation (DA) methods where the tangent or adjoint models of the observation operators, the RTMs in this case, are needed, ensemble DA methods, such as the ensemble Kalman filter (EnKF; Evensen 1994) is much easier to implement when the observation operators are complex. EnKF derives flow-dependent background error covariances from ensemble forecasts, and can thereby update many state variables not directly observed; therefore, it has additional benefit for DA applications at the convective scale where observed parameters tend to be more limited compared to the number of state variables (Snyder and Zhang 2003; Tong and Xue 2005; Aksoy et al. 2009; Dowell et al. 2011; Yussouf et al. 2015). Considering these aspects, most studies on the assimilation of all-sky geostationary IR observations into convective-scale models have chosen to use EnKF or its variants, either with simulated observations (e.g., Cintineo et al. 2016; Zhang et al. 2016; Minamide and Zhang 2017) or real observations (e.g., Honda et al. 2018; Jones et al. 2018; Minamide and Zhang 2018; Zhang et al. 2018; Sawada et al. 2019; F. Zhang et al. 2019; Y. Zhang et al. 2019; Jones et al. 2020).

Most earlier attempts to assimilate ABI radiance data used synthetic or simulated observations via observing system simulation experiments (OSSEs). For example, Otkin (2010, 2012a) assimilated hourly clear and cloudy synthetic ABI window channel brightness temperatures (BTs) onto a 12-km grid spacing model using ensemble adjustment Kalman filter (EAKF; Anderson 2001) implemented in the Data Assimilation Research Testbed (DART) system for an extratropical cyclone case, and examined the impact of different horizontal localization radii. Otkin found assimilating window channel BTs improves cloud analyses and forecasts but also has the tendency to degrade moisture and thermodynamic fields unless using a small localization radius of 100 km. Otkin (2012b) assimilated three ABI water vapor channels and demonstrated their enhanced abilities to improve not only cloud analysis but also moisture, temperature, and wind fields, as compared to the window channel. In the OSSE study of Jones et al. (2013b), simulated ABI 6.95-μm water vapor channel BTs were assimilated in a more frequent 5-min assimilation intervals together with simulated WSR-88D radar reflectivity and radial winds at a 15-km grid spacing using EAKF for a cool-season extratropical cyclone. They found that ABI BTs can provide additional improvements in the mid and upper tropospheric cloud and humidity analyses when assimilated together with radar observations. The follow-on study of Jones et al. (2014) examined the impacts of the ABI BT and radar DA of Jones et al. (2013b) on 1–3-h forecasts and found the assimilation of ABI 6.95-μm BTs primarily improved water vapor and ice cloud forecasts whereas radar observations improved the lower and middle tropospheric hydrometeor fields most. The effects of assimilating BT data decreased rapidly after 1 h of forecast. Compared to the above studies that used non-convection-allowing resolutions, Cintineo et al. (2016) assimilated bias-corrected synthetic ABI BTs and radar observations into a 4-km grid spacing convective-allowing model. Different covariance localization radii of ABI BTs were tested, and a radius of 28 km was found to perform best. All these tests were done with simulated ABI observations through OSSEs.

Following those previous OSSE efforts, convective-scale assimilations of real all-sky *Himawari-8/9* AHI and GOES-R ABI IR BTs have been carried more recently. Using the local ensemble transform Kalman Filter (LETKF; Hunt et al. 2007), Honda et al. (2018) assimilated all-sky AHI BTs from *Himawari-8* for the first time to investigate its impact on the analyses and forecasts of Supertyphoon Soudelor (2015) and found the analyzed TC structure and intensity forecasts are improved. Using the Pennsylvania State University (PSU) WRF-EnKF system, Minamide and Zhang (2018) demonstrated that hourly assimilation of AHI BTs can constrain Typhoon Soudelor's inner-core moist convection and develop

a more resilient initial vortex. Again using the PSU EnKF system, Zhang et al. (2018) presented successful assimilation of real GOES-R ABI observations for a tornadic thunderstorm case and found all-sky ABI radiances can help suppress spurious model clouds before storm initiation and improve the prediction of midlevel mesocyclone and low-level vortex signatures. Sawada et al. (2019) assimilated multichannel AHI BTs into a 2-km limited area model in Japan region using LETKF with a running-in-place (RIP) technique that assimilated the same observations a number of times. The results indicated positive contribution of AHI infrared radiances to the forecast of isolated and disorganized convective activities in much smaller spatiotemporal scales. Following Zhang et al. (2018), real ABI IR radiances were assimilated together with radar data and additional benefit of the former was achieved. Jones et al. (2020) showed that all-sky radiance assimilation improved convective initiation forecast of severe storms in several instances within the Warn-on-Forecast System (Wheatley et al. 2015).

To better utilize all-sky IR radiances, Minamide and Zhang (2017) proposed, for OSSE experiments for a tropical cyclone (TC), an adaptive observation error inflation (AOEI) method aiming at limiting erroneously large analysis increments when the observation innovation is large. In Minamide and Zhang (2019), an adaptive background error inflation (ABEI) method was further proposed to help initiate convection within TCs when the background spread is too small. Otkin et al. (2018) and Otkin and Potthast (2019) developed and tested a nonlinear bias correction algorithm for IR BT data and obtained promising results.

Despite the afore-referenced work, the number of studies assimilating real all-sky GOES-R ABI data into convection-allowing/resolving models for improving short-range convective-scale weather prediction is still very limited. There are still many issues needing more careful investigation before such all-sky observations can be assimilated routinely or operationally to achieve robust positive impacts. Within the United States, it is very desirable to test the data impact within a DA system, such as the Gridpoint Statistical Interpolation (GSI; Kleist et al. 2009) system, that is used by operational forecast systems for easier operational implementation. Also, a better understanding is needed of how the assimilation of cloudy and clear-sky IR BT observations impacts, both positively and negatively, the analyzed storms as well as their environment in terms of individual state variables and on subsequent forecasts. Further, for high-density all-sky BT observations, appropriate data thinning and covariance localization need further testing. Finally, how to better handle highly nonlinear observation operator in EnKF also deserves attention.

In this study, we explore the impact of assimilating GOES-R ABI IR BTs on improving the analysis and forecast of a mesoscale convective system (MCS) case in a 3-km convection-allowing model using GSI-based EnKF system. The ability of real all-sky ABI IR observations to build up observed storms and remove spurious model storms are examined. We also compare the results of using different ABI horizontal localization radii and assimilation intervals, as well as using two variations of the EnKF ensemble mean update equation. Significant differences are found with the assimilation results between two different ways of calculating the ensemble mean of BT observation prior, due to the high nonlinearity of the BT observation operator in the presence of clouds. Furthermore, the background error correlations between simulated BTs and model state variables, and the corresponding analysis increments are analyzed to gain understanding on how BT data impact different state variables. It is worth noting that certain conclusions drawn based on a single test case in this study will need to be tested with more cases for them to be robust. The use of a single case makes it easier to investigate specific behaviors in more details. We believe the results on the ways the ensemble mean of ABI observation priors are calculated are general.

The rest of this paper is organized as follows. Section 2 provides a brief description of the GOES-R ABI observations used in this study. Descriptions of the EnKF algorithms, observation operator, and experiment design are given in section 3. Results are given in section 4. Summary and conclusions are provided in section 5.

2. *GOES-16* ABI observations

GOES-16, launched by NASA and the National Oceanic and Atmospheric Administration (NOAA) on 19 November 2016, is the first geostationary weather satellite of the NOAA GOES-R series. ABI is a multichannel passive imaging radiometer on GOES-R series designed to sense various surface parameters and atmospheric phenomena. It has 16 spectral bands in total, including 2 visible channels, 4 near-infrared channels and 10 infrared channels (Schmit et al. 2017). The spatial resolution of each ABI channel is between 0.5 and 2 km at subsatellite point, which is much higher than its predecessors. Three water vapor channels (channels 8–10) strongly affected by atmospheric humidity are located at the middle and upper troposphere. To avoid interchannel observation error correlations caused by the overlapping measurements of adjacent channels, only observations from ABI channel 10 (7.3 μm) are assimilated, as it is more sensitive to low level moisture and cloud properties.

The ABI BT data used in this study are the 5-min-interval Multiband Cloud and Moisture Imagery products (MCMIP) over a CONUS domain downloaded from NOAA's Comprehensive Large Array-Data Stewardship System (CLASS). These ABI brightness temperature products are derived from the calibrated ABI level 1b radiances. Satellite zenith angle, which will be an input parameter to the radiative transfer model used as the ABI observation operator, is calculated following C. Yang et al. (2017), along with geolocation for each BT pixel. Also, a parallax correction (e.g., Wang and Huang 2014) using ABI L2 Cloud Top Height (ACHA) products is applied for cloudy BTs to reduce the "parallax error" caused by the slantwise nature of satellite observation.

Following Zhang et al. (2018) and (Y. Zhang et al. 2019), no extra thinning and quality control procedures are applied to BT observations in this study so that detailed structures of convective-scale weather systems can be better captured by the high-resolution satellite observations. Although spatial

error correlations may exist with the high-resolution observations, their effects can be partly reduced by inflating observation error somewhat. We do not apply bias correction here because the rapid change in bias characteristics in both model and observations at the convective scale makes online bias correction (e.g., Zhu et al. 2019) more difficult. Bias correction of cloudy radiance data for convective scale DA is an important research topic that is beyond the scope of this study.

3. Methodology and experimental setup

a. GSI-based EnKF and radiative transfer model

An enhanced GSI-based EnKF DA system is used in this study for the assimilation of all-sky GOES-R ABI BTs. The ensemble DA algorithm used is the ensemble square root filter (EnSRF; Whitaker and Hamill 2002).

EnSRF is a serial filter, which means that the model state is updated by one observation at a time. Within GSI EnKF, this update process is done using the so-called "scalable" implementation of the filter (Anderson and Collins 2007), in which both model state variables and precalculated observation priors are updated within the observation loop so as to avoid recomputing the observation priors after each observation is assimilated. The "scalable" implementation yields theoretically the same results as the observation recomputing method only when the observation operator is linear.

According to Whitaker and Hamill (2002), by separating the prior ensembles into an ensemble mean and perturbations, model variables updated by a single observation follows the equations below in EnSRF:

$$\overline{\mathbf{x}}^a \equiv \overline{\mathbf{x}}^b + \mathbf{K}_x(\mathbf{y}^o - \overline{\mathbf{y}}^b), \tag{1}$$

$$\mathbf{x}_i'^a = \mathbf{x}_i'^b - \tilde{\mathbf{K}}_x\mathbf{y}_i'^b, \tag{2}$$

where $\overline{\mathbf{x}}$ and $\mathbf{x}_i'$ are, respectively, the ensemble mean and perturbation of state vector $\mathbf{x}$, with superscripts a denoting the analysis and b denoting the background. Subscript i indicates the ith ensemble member. The term $\mathbf{y}^o$ is the (single) observation being assimilated; $\overline{\mathbf{y}}^b$ and $\mathbf{y}_i'^b$ indicate the mean and perturbation of observation prior updated by the previous observations. The term $\mathbf{K}_x$ is the Kalman gain and $\tilde{\mathbf{K}}_x$ the "reduced" Kalman gain to be given later.

After the update of model state variables, the ensemble mean and perturbation priors of the observation being assimilated are also updated before the assimilation of the next observation. Similar to Eqs. (1) and (2), the corresponding equations are

$$\overline{\mathbf{y}}^a = \overline{\mathbf{y}}^b + \mathbf{K}_y(\mathbf{y}^o - \overline{\mathbf{y}}b), \tag{3}$$

$$\mathbf{y}'^a_i = \mathbf{y}'^b_i - \tilde{\mathbf{K}}_y\mathbf{y}'^b_i. \tag{4}$$

Here $\overline{\mathbf{y}}^a$ and $\mathbf{y}'^a_i$ are the updated observation prior mean and perturbations after assimilating the current observation.

In Eqs. (1) and (3), $\mathbf{K}_x$ and $\mathbf{K}_y$ represent the Kalman gains defined by Eqs. (5) and (6) below, and are used to update the mean state variables and observation priors, respectively. The $\tilde{\mathbf{K}}_x$ and $\tilde{\mathbf{K}}_y$ terms in Eqs. (2) and (4) are the "reduced" Kalman gains used to update the perturbations and are given in Eqs. (7) and (8):

$$\mathbf{K}_x = \mathbf{P}\mathbf{H}^{\mathrm{T}}(\mathbf{H}\mathbf{P}^f\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}, \tag{5}$$

$$\mathbf{K}_y = (\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}})(\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}, \tag{6}$$

$$\tilde{\mathbf{K}}_x = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{R}}}\right)^{-1}\mathbf{K}_x, \tag{7}$$

$$\tilde{\mathbf{K}}_y = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}\mathbf{P}\mathbf{H}^{\mathrm{T}} + \mathbf{R}}}\right)^{-1}\mathbf{K}_y, \tag{8}$$

where $\mathbf{P}$ is the background-error covariance matrix and $\mathbf{R}$ is the observation-error covariance matrix. The algorithms in Eqs. (1)–(4) are repeated for all observations, one at a time in a serial manner.

To use the parallel (Anderson and Collins 2007) algorithm described above, the mean and perturbations of all observation priors need to be computed first by applying the forward observation operator to the background state before actual assimilation. In GSI, the observation prior mean $\overline{\mathbf{y}}$ and perturbation $\mathbf{y}'$ vectors are calculated by

$$\overline{\mathbf{y}} = H(\overline{\mathbf{x}}^b), \tag{9}$$

$$\mathbf{y}'_i = H(\mathbf{x}_i) - \overline{H(\mathbf{x}^b)}, \tag{10}$$

where H is the observation operator.

Note that for the calculation of $\overline{\mathbf{y}}$ in Eq. (9), the observation operator is applied to the mean background state as in current GSI EnKF system. Here, we will test an alternative that calculates $\overline{\mathbf{y}}$ as the mean of the observation priors of all members, i.e., as $\overline{H(\mathbf{x}^b)}$. The two are the same if the observation operator is linear. For nonlinear operators, such as the radiative transfer model in cloudy regions, the choice of $H(\overline{\mathbf{x}}^b)$ or $\overline{H(\mathbf{x}^b)}$ can make big difference, as will be shown later. We note here that this alternative formulation is actually commonly used in previously published EnKF algorithms (e.g., Houtekamer and Mitchell 2001; Evensen 2009).

For the assimilation of satellite radiance including the ABI BT in this paper, the Community Radiative Transfer Model (CRTM; Han et al. 2006) is used as the forward observation operator. CRTM is a rapid radiative transfer model developed by Joint Center for Satellite Data Assimilation (JCSDA) and it is integrated in the GSI EnKF system. This study uses CRTM version 2.3.0 to compute the observation priors of ABI infrared BT in clear and cloudy regions. For cloudy BT, mixing ratios of hydrometeors including cloud water (q_c), cloud ice (q_i), snow (q_s), rainwater (q_r), and graupel (q_g) are part of the input, in addition to temperature, water vapor, 10 m winds, and land surface parameters. The effective radii of hydrometeors needed by CRTM are calculated based on the drop size distribution assumptions of the Thompson microphysics scheme (Thompson et al. 2008).

TABLE 1. Descriptions of experiment settings.

Experiment name	Observation prior mean ($\overline{\mathbf{y}}$)	EnKF interval (min)	Horizontal localization radius (km)	Vertical localization radius (scale height)
CTRL	—	—	—	—
DETER	—	—	—	—
MHX_15km	$\overline{H(\mathbf{x}^f)}$	5	15	4
MHX_50km	$\overline{H(\mathbf{x}^f)}$	5	50	4
HMX	$H(\overline{\mathbf{x}}^f)$	5	30	4
MHX	$\overline{H(\mathbf{x}^f)}$	5	30	4
MHX_15min	$\overline{H(\mathbf{x}^f)}$	15	30	4

b. *Experimental setups*

The MCS event from 12 to 13 July 2018 over the U.S. Central Plains is selected as a test case for examining the impacts of assimilating *GOES-16* ABI BT data. The Weather Research and Forecasting Model with Advanced Research dynamic core (WRF-ARW) version 3.8.1 (Skamarock et al. 2008) is used as the forecast model.

Our experiments use a single domain with a 3-km horizontal grid spacing covering the central Great Plains, and there are 300 × 300 grid points in the horizontal. There are 50 vertical layers with vertical grid stretching and the model top is at 50 hPa. The EnKF system includes 40 WRF ensemble forecast members with different combinations of surface layer and planetary boundary layer (PBL) parameterization schemes following Kong et al. (2020).

The initial conditions of the ensemble forecasts are created by adding perturbations derived from the 3-h forecasts of 1500 UTC cycle operational Short-Range Ensemble Forecast (SREF) system to the operational North American Mesoscale Forecast System (NAM) analysis valid at 1800 UTC (see Labriola et al. 2021 for more details). Perturbations are constructed by taking the positive and negative differences between pairs of members from the 24 SREF members. Some of these differences are reused to generate a total of 39 perturbations. The perturbations of u and v wind components, temperature, and water vapor mixing ratio are then scaled with domain-averaged standard deviations of 2 m s^{-1}, 1°C, and 0.5 g kg^{-1}, respectively, before being added to the NAM analysis. The 40 ensemble members are created with those 39 perturbed members (i.e., member 2–40) and one unperturbed member (member 1). The 3-h spinup ensemble forecasts are carried out from 1800 to 2100 UTC 12 July when the assimilation of ABI radiance starts.

For model physical parameterizations, we use partially double-moment Thompson microphysics scheme (Thompson et al. 2008), the Rapid Radiative Transfer Model for General Circulation Models (RRTMG) for longwave and shortwave radiation (Iacono et al. 2008) and the Noah land surface model (Chen and Dudhia 2001). The surface layer schemes used for the ensemble members include the revised MM5 Monin–Obukhov scheme (Jiménez and Dudhia 2012), the Monin–Obukhov Janjić Eta scheme (Janjić 1996) and that from the Mellor–Yamada–Nakanishi–Niino (MYNN) PBL scheme (Nakanishi 2001). The PBL schemes used are the Yonsei University (YSU) scheme (Hong et al. 2006), Mellor–Yamada–Janjić (MYJ) scheme (Janjić 1994), the MYNN level 2.5 scheme (Nakanishi and Niino 2006), and the Asymmetrical Convective Model 2 (ACM2) scheme (Pleim 2007). Using different combinations of PBL and surface layer schemes aims to account for model error, and helps increase the ensemble spread. This configuration is similar to that of the 2018 Hazardous Weather Testbed (HWT) Spring Experiment employed at the Center for the Analysis and Prediction of Storms (CAPS). Cumulus parameterization is not used in our 3-km domain.

Several experiments are performed. The control experiment labeled as "CTRL" assimilates no observations within the 2100–2300 UTC time window when other DA experiments assimilate radiance data. CTRL continues the 2-h long ensemble forecasts with the same configuration until 2300 UTC, when a deterministic forecast is started as in other DA experiments, except from the ensemble mean of forecasts. This CTRL experiment serves as a no-DA reference for the DA experiments to compare against.

An alternative to the no-DA CTRL used here as reference is a deterministic forecast starting from 1800 UTC NAM analysis that is run all the way through the end of the forecast period of DA experiments. Both forecasts can be used as a baseline for comparison purpose, although the current CTRL has procedurally smaller differences from the DA experiments which involve ensemble forecasts through 2300 UTC. The convection in the ensemble mean initial condition at 2300 UTC for the deterministic forecast in the current CTRL is subject to smoothing due to ensemble averaging, which may lead to weaker convection in the initial forecast. The deterministic forecast from 1800 UTC would not suffer from such smoothing from 2300 UTC. This alternative forecast is also performed and is named DETER in Table 1. Its forecast results will be included for comparison.

After 2 h of cycled DA, 4-h deterministic forecasts are run from the ensemble mean final analysis at 2300 UTC for each experiment, with NAM analyses at 3-h intervals providing the lateral boundary conditions.

The first DA experiment (referred to as "MHX") assimilates all-sky ABI channel 10 BTs every 5 min using the modified EnSRF algorithm in GSI mentioned in section 3a that the observation prior mean is calculated from $\overline{H(\mathbf{x}^b)}$, while the second DA experiment (referred to as "HMX") uses $H(\overline{\mathbf{x}}^b)$ to calculate observation prior mean. The third experiment is similar to "MHX," except that it assimilates ABI BTs every 15 min (referred to as "MHX_15min"). These experiments use a horizontal localization radius of 30 km. More details can be found in Table 1.

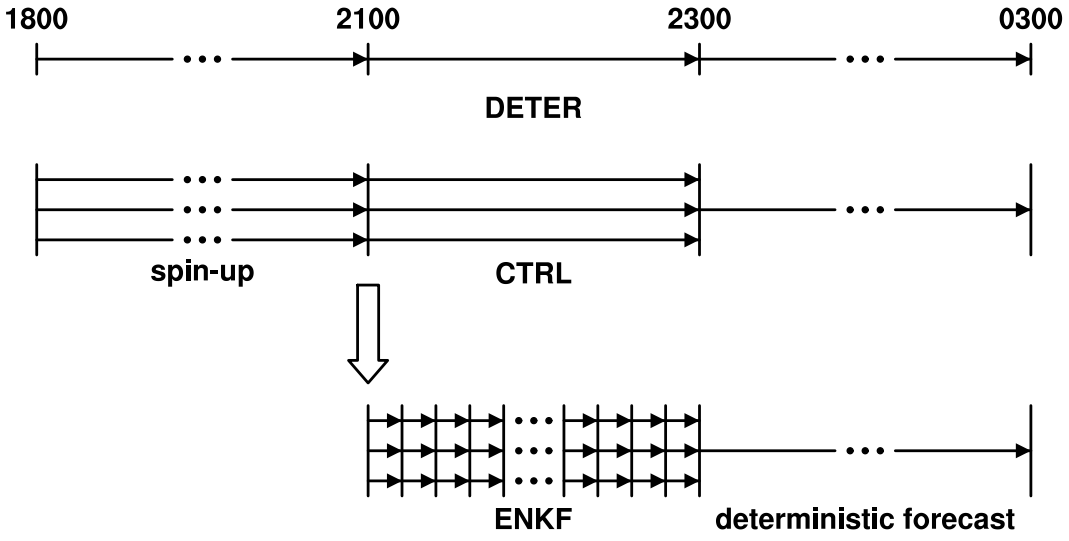

FIG. 1. Flow diagram of experiment design for DETER, CTRL, and cycling EnKF experiments. The 3-h spinup ensemble forecasts with 40 members are initialized from 1800 UTC operational NAM analysis plus perturbations derived from 3-h SREF forecasts. Deterministic forecasts are carried from 2300 to 0300 UTC. ABI BT data are assimilated from 2100 through 2300 UTC at 5- or 15-min intervals. CTRL contains ensemble forecasts without DA through 2300 UTC when a deterministic forecast starts from its ensemble mean. DETER is a deterministic forecast from 1800 to 0300 UTC initialized from the 1800 UTC NAM analysis.

Two additional sensitivity experiments named as "MHX_15km" and "MHX_50km" are carried out to test different localization radii for all-sky ABI BTs. These two experiments are the same as MHX except that the horizontal localization radius is 15 or 50 km, respectively. Figure 1 shows the experiment workflows.

In all DA experiments, the observation height of ABI radiance is assigned to the level where the atmospheric transmittance changes the fastest, i.e., at the peak level of simulated weighting function in clear-sky region or the ensemble mean background cloud top when cloud is present in the column. The height for ABI observation is only used for vertical localization purpose within the EnSRF algorithm. For clear-sky ABI observations, observation error is set to 3 K, same as that used in Zhang et al. (2018). An AOEI technique (Minamide and Zhang 2017) is applied to adaptively inflate observation error variance when the magnitude of calculated observation innovation is large. Spatial covariance localization is applied with the Gaspari and Cohn (1999) method with a horizontal localization radius ranging between 15 and 50 km, as described earlier. A relative broad vertical localization radius of 4.0 in units of scale height is used in this study following Jones et al. (2018). To help maintain ensemble spread and prevent the filter divergence, the relaxation to prior spread (RTPS; Whitaker and Hamill 2012) inflation method with an inflation parameter of 0.95 as used in Sawada et al.(2019) and Kong et al.(2020) is applied in each DA cycle. We also tested other covariance inflation methods, including the relaxation to prior perturbation (RTPP; Zhang et al. 2004) method and a combination of RTPS and RTPP. The observation priors and posteriors had the smallest biases and root-mean-square innovations (RMSIs) with RTPS. Meanwhile, the ensemble spreads of RTPS were better maintained compared to using RTPP or using RTPS and RTPP together (figures not shown). The atmospheric state variables analyzed by the EnSRF include three wind components, potential temperature, water vapor, hydrometeor mixing ratios, and the number concentrations of ice and rainwater. For hydrometeors and number concentrations, an ad hoc adjustment used in Zhang et al. (2016) was applied to the posterior ensembles to avoid negative water vapor and hydrometeors; certain positive values are reduced when negative values are set to zero to keep the ensemble mean unchanged.

4. Results

a. *Sensitivity experiments to horizontal localization radius*

Figure 2 shows the time series of the domain-averaged biases and RMSIs for clear and cloudy ABI channel-10 BTs during the entire DA cycles in experiments MHX, MHX_15km, MHX_50km, when different horizontal localization radii of 30, 15, and 50 km are used, respectively. Here, the bias and root-mean-squared innovation (RMSI) are calculated as

$$\text{bias} = \frac{1}{M}\sum_{m=1}^{M}[\mathbf{y}_m - H(\overline{\mathbf{x}}_m)], \tag{11}$$

$$\text{RMSI} = \sqrt{\frac{1}{M}\sum_{m=1}^{M}[\mathbf{y}_m - H(\overline{\mathbf{x}}_m)]^2}, \tag{12}$$

where M is the number of clear or cloudy observations within the entire WRF domain. Note that the innovation used for the calculating bias and RMSI is defined in terms of the

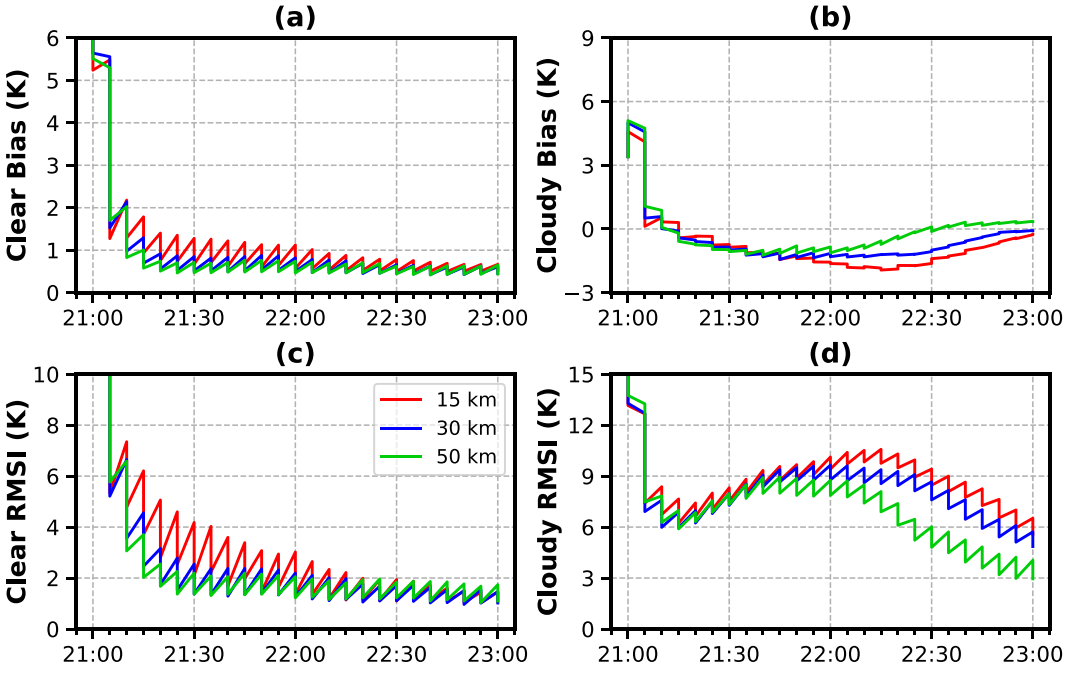

FIG. 2. Time series of (a),(c) clear-sky and (b),(d) cloudy-sky channel-10 (top) BT biases [$\mathbf{y} - H(\overline{\mathbf{x}})$] and (bottom) RMSIs for the priors or background forecasts (typically high values at a given time) and posteriors or analyses (typically low values at a given time) during the entire EnKF cycles, for experiments with 15- (red), 30- (blue), and 50-km (green) localization radius.

observation minus the ensemble mean prior [$\mathbf{y}^o - H(\overline{\mathbf{x}}^b)$] or posterior [$\mathbf{y}^o - H(\overline{\mathbf{x}}^a)$]. The channel-10 observations to be assimilated at locations where ABI channel-14 (11.2 μm) BTs are lower (higher) than 285 K are treated as cloudy (clear) sky observations. This criterion is similar to that used in Zhang et al. (2016, 2018). Under clear sky, within the first 75 min the biases and RMSIs are the largest using the 15-km radius (Figs. 2a,c). Because most of the misfit to observations comes from the background having spurious clouds or convection, the results indicate that the effectiveness in removing spurious convection is reduced when using a small 15-km localization radius. The differences are relatively small between 30- and 50-km radii. However, the prior RMSIs for the 50-km radius case became the largest in the last 30 min of DA, apparently because of the generation of some new small spurious storms during that period. The differences in RMSI are relatively small though.

In cloudy regions, the 15-km radius case still has the overall largest RMSI, especially during the second hour of DA. The 50-km case has the smallest bias and RMSI (Figs. 2b,d). Negative cloudy biases occur in these experiments after 2115 UTC with increase in RMSIs. Unlike the clear sky regions, the differences in cloudy biases and RMSIs among the three experiments are small in the earlier cycles but become larger after 2140 UTC. As will be discussed more later, an important role of cloudy radiance DA is to build up clouds and convection that are missing in the forecast background. The 30- and 50-km radii appear to better able to build up observed convection missing in the background, presumably due to increased impact of the observations when they influence more grid points in the EnKF updating.

Neighborhood equitable threat scores (ETS; Mason 2003) of composite reflectivity forecasts are calculated and shown in Fig. 3. A 15-km neighborhood radius is used to accommodate small displacement errors (Ebert 2009). For the 20-dBZ threshold (Fig. 3a), the 30-km localization radius has the highest ETS, while the 50-km radius scores the lowest almost throughout the forecast. For the 40-dBZ threshold, the 50-km radius still underperforms the other two. The 15- and 30-km radii alternate to have higher scores several times during the 4 h of forecasts, and have similar performance on average. Overall, the 50-km radius clearly has the lowest ETSs despite the smallest observation minus background (O-B) and observation minus analysis (O-A) biases and RMSIs during the DA cycles. Smaller O-A differences do not necessarily mean more accurate analyses because more state variables are not directly measured and do not directly contribute to the O-A calculation. The 5-min forecasts may be too short to gauge the quality of analyses in terms of O-B statistics. In the absence of truth about the state variables, ensuing forecasts often have to be relied upon to gauge the true quality of analysis. The poorer forecast in MHX_50km is most likely due to the location displacement between observed and analyzed convection, as well as the generally less accurate moisture and thermodynamic analyses resulting from overspreading of cloudy radiance observation information when using a relatively large radius of influence (figures not shown).

The results of above sensitivity experiments suggest that a horizontal localization radius of 30 km is overall the best

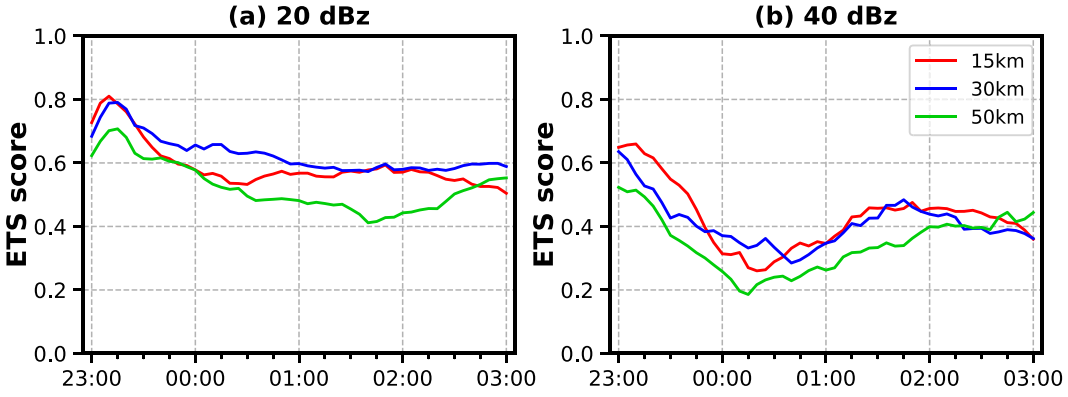

FIG. 3. Neighborhood ETSs of composite reflectivity forecasts calculated with a 15-km neighborhood radius at (a) 20-and (b) 40-dBZ thresholds for experiments using 15- (red), 30- (blue), and 50-km (green) localization radii.

choice among the three tested, and it is, therefore, used in all remaining DA experiments in this study. This localization radius is similar to that used in the previous works for mesoscale convective system applications at convection-allowing resolutions (e.g., 28 km in Cintineo et al. 2016; 30 km in Zhang et al. 2018; 35 km in Otkin and Potthast 2019; and 36 km in Jones et al. 2020) when assimilating ABI or SEVIRI data. It is, however, larger than the radii typically used for radar DA on ~3-km grids (e.g., Sobash and Stensrud 2013; Wheatley et al. 2015; Kong et al. 2018; Tong et al. 2020).

b. *Impact of all-sky ABI radiance on EnKF analyses*

Comparison between observed and simulated ABI channel 10 BTs for experiments CTRL and MHX at 2100 (the first DA cycle), 2200, and 2300 UTC (the last DA cycle) is shown in Fig. 4. In the ensemble prior mean at 2100 UTC, there is a large region of spurious clouds from southeastern Minnesota to western Wisconsin (Fig. 4b), where the sky is observed to be mostly cloud free (Fig. 4a). After the first EnKF cycle, much of the cloud in this region is significantly reduced by the assimilation of ABI radiance (Fig. 4c). Spurious background clouds in northern Nebraska are also much reduced. By 2200 UTC, after a full hour of ABI radiance assimilation at 5-min intervals, almost all spurious clouds can be effectively cleared (Fig. 4f) even though the forecast background still has significant overprediction (Fig. 4e). The clouds in Nebraska are very well reproduced in the analysis while those in eastern South Dakota and southern Minnesota are somewhat underestimated in size and intensity (Fig. 4f). After 2 h of DA, the difference between the analyzed (Fig. 4i) and observed (Fig. 4g) BTs becomes quite small, indicating that the cycled EnKF DA is very effective in ensuring close fit of the analysis to observations (Fig. 4i). As a comparison, all the major systems found in the 2100 UTC background (Fig. 4b) continue to develop in CTRL within the 2-h free ensemble forecasts with much wider coverage of clouds but lower intensities than observations (not shown). We point out here that even though the analyzed BT at 2300 UTC is very close to observed values (Fig. 4i), the ensemble forecast mean in CTRL at the same time still significantly overforecast the cloud coverage (Fig. 4h), suggesting that errors in the analyzed state variables may still contain significant error, causing quite fast forecast error growth.

The time series of domain-averaged bias, RMSI, ensemble spread, and prior consistency ratio (CR) (Dowell et al. 2004; Aksoy et al. 2009; Dowell and Wicker 2009) during each EnKF cycle are also calculated for all three EnKF experiments ("HMX," "MHX," and "MHX_15min"). The ensemble spread and CR are defined as

$$\text{spread} = \sqrt{\frac{1}{M}\sum_{m=1}^{M}\left\{\frac{1}{N-1}\sum_{n=1}^{N}[H(\mathbf{x}_{m,n}) - \overline{H(\mathbf{x}_{m,n})}]^2\right\}}, \quad (13)$$

$$\text{CR} = \frac{\sigma_{\text{obs}}^2 + \text{spread}^2}{\text{RMSI}^2}, \quad (14)$$

where $\mathbf{y} - H(\overline{\mathbf{x}})$ represents the prior or posterior innovation and σ_{obs} is the observation error standard deviation. The terms M and N are the number of observations and ensembles, respectively. Note that the quantification of the observation error for satellite radiance is difficult, we choose to set σ_{obs} to 3 K (the clear sky value without inflation) here, which is different from the adaptively inflated observation error (Minamide and Zhang 2017) used in the EnKF.

As shown in Fig. 5, assimilating all-sky ABI radiance has the greatest impact in the first few cycles as biases for all three experiments significantly decrease from their initial values of ~6.9 K, and biases remain relatively small thereafter (Fig. 5a). Bias in MHX is closest to 0 K, while in MHX_15min, the prior or forecast biases are much larger due to the longer DA cycles used (15 vs 5 min) that results in more forecast errors. In experiment HMX, the BT bias is reduced to about −1 K after the first cycle and remains negative during the entire EnKF cycles. This negative bias is associated with excessive background water vapor removal over cloudy regions in the background when $H(\overline{\mathbf{x}})$ is used to derive the prior innovation, as drier moisture profiles lead to higher BTs. Meanwhile, more spurious clouds are cleared in experiment HMX, which further reduces positive biases.

RMSIs also decrease quickly in the first 15 min, owing to effective removal of spurious clouds. The reduction is largest in

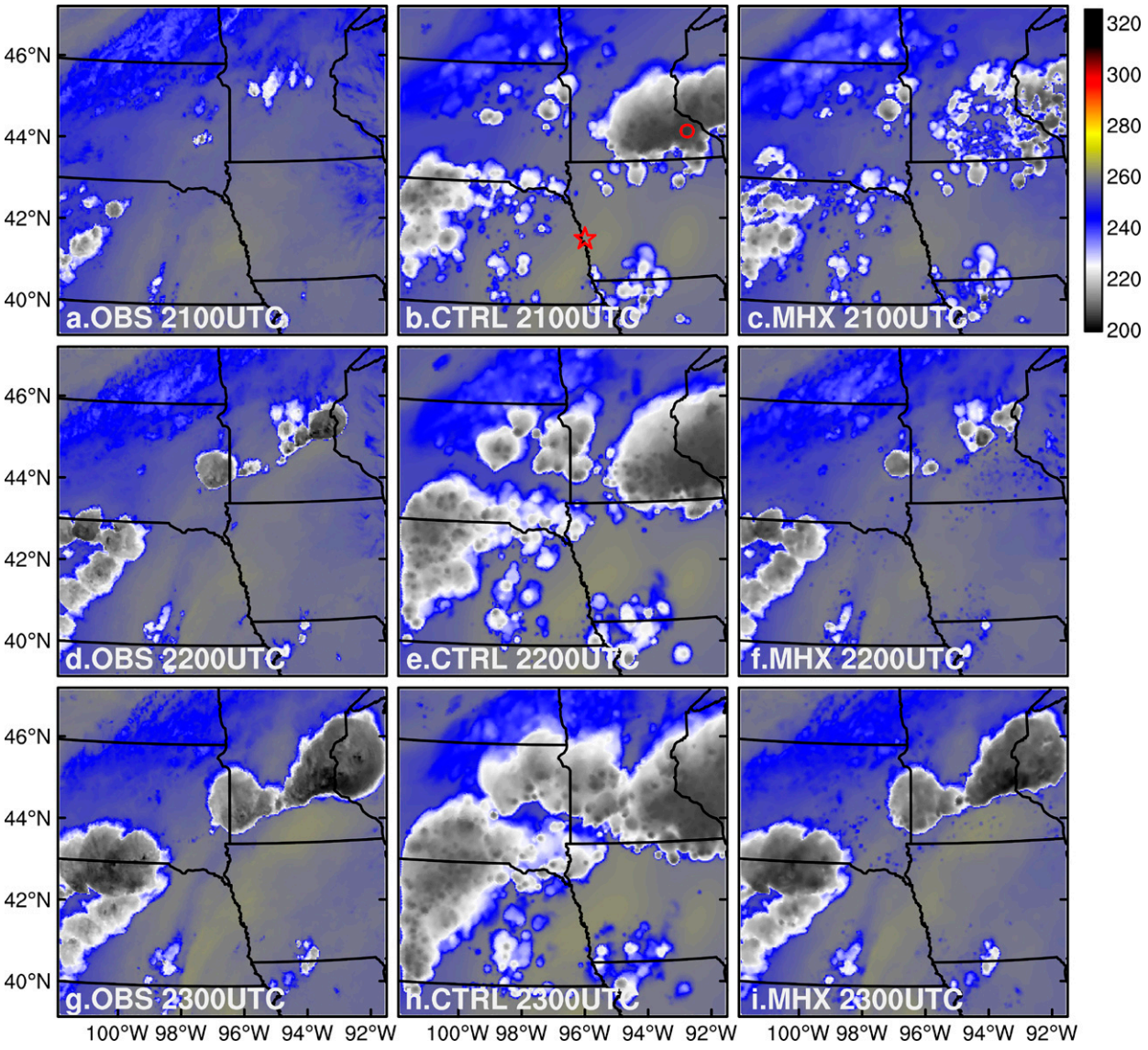

FIG. 4. (left) Observations and simulated brightness temperatures (K) of ABI channel 10 for (center) CTRL ensemble mean forecasts and (right) ensemble mean analyses of experiment MHX valid at (a)–(c) 2100, (d)–(f) 2200, and (g)–(i) 2300 UTC. The red star and circle in (b) denote clear and cloudy pixels, respectively. Figure shows the entire WRF domain.

the first cycle (Fig. 5b). However, there is gradual growth in RMSI from 2115 to 2200 UTC, resulting from a slower development of new clouds in the analyses and forecasts compared to observations. The RMSIs are decreased somewhat again after 2200 UTC and reach minimum at the end of DA cycles. The reduction of RMSI in HMX is initially larger than that in MHX but becomes very close to that of MHX until 2200 UTC when the RMSI of MHX becomes the smallest. At the end of all DA cycles, the analysis RMSIs are reduced to about 3.1, 3.6, and 5 K for MHX, HMX, and MHX_15min, respectively. A larger RMSI of MHX_15min is due to larger growth of RMSI during the longer 15-min forecasts, and fewer DA cycles that serve to reduce the innovations (Fig. 5b).

The ensemble spreads are similar after 20 min for HMX and MHX. The ensemble spread has been maintained at this level using the RTPS covariance inflation; however, they are still significantly lower than the corresponding RMSI. The RMSIs of MHX_15min are more than twice as large, due to the assimilation of much fewer observations. Not surprisingly, prior CRs for all experiments are consistently below the optimal value of approximately 1, suggesting that the ensemble spread is low and/or observation error is specified to be too small [see Eq. (14)]. We do point out that the observation error variance σ_{obs} used in Eq. (14) is 3 K while within EnKF it is inflated to much larger values within cloudy regions that have large background innovations. This would significantly increase the actual CR in these areas.

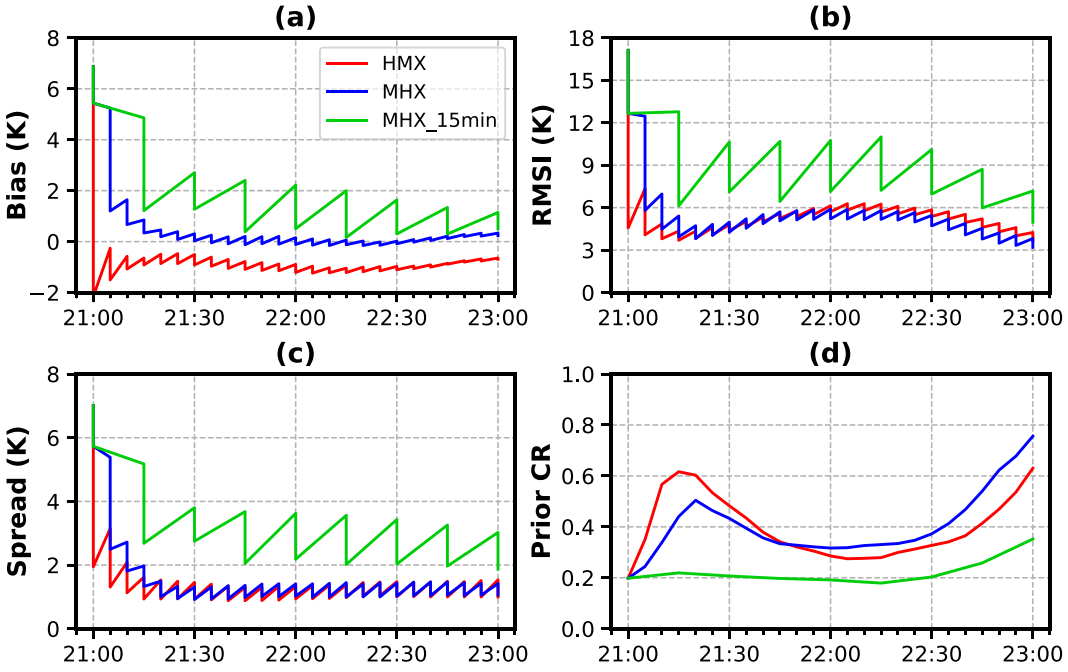

FIG. 5. Time series of channel 10 BT (a) biases [$\mathbf{y} - H(\overline{\mathbf{x}})$], (b) RMSIs, (c) ensemble spreads, and (d) prior consistency ratios during the entire EnKF cycles, for experiments HMX (red), MHX (blue), and MHX_15min (green).

The CRs do increase in all three experiments in the last 45 min with that of MHX reaching 0.75 at the end. The underdispersion (too low spread) problem is regularly seen in high-resolution radar DA studies; the horizontal resolutions of radar data are similar to that of BTs in this study (Aksoy et al. 2009; Dowell and Wicker 2009; Dowell et al. 2011; Snook et al. 2011). Dense observations tend to significantly decrease ensemble spread, as seen in Fig. 5c. We tried adding random additive perturbations to the model state variables, but the results were not much improved; they are, therefore, not employed in experiments shown here.

For independent verification, simulated composite radar reflectivity fields from the final ensemble mean analyses are compared with observed one in Fig. 6. All DA experiments are able to capture the more intense precipitation band (Figs. 6d–f) while the no-DA CTRL shows a much broader band of weak precipitation and there is also significant displacement error (Fig. 6b). For CTRL it should be noted that the reflectivity at this time (Fig. 6b) is the mean of 5-h ensemble forecasts (cf. Fig. 1); therefore, it tends to underestimate the peak values of reflectivity and overestimate the spatial coverage than individual ensemble members due to smoothing across the members. As a comparison, the reflectivity at 2300 UTC in DETER is noticeably stronger with more distinct convective cores and is also less widespread than in CTRL. However, the convective line in southern Minnesota is significantly displaced to the south and the north–south-oriented line in western Nebraska is too strong and displaced to the west (Fig. 6c). In experiments MHX and MHX_15min, the reflectivity fields generally match the observation much better. The higher (>35 dBZ) reflectivity values show a better linear structure and at locations close to observed, although their spatial coverages are slightly larger than observed. In MHX_15min, more spurious storms, in terms of high reflectivity, are analyzed than in the other two DA experiments. Overall, the analyzed composite reflectivity is much close to observations than those in CTRL and DETER, which assimilate no BT data. In section 4c, we compare forecasts starting from the final analyses at 2300 UTC.

c. *Impact of BT data assimilation on forecasts*

Last subsection showed benefits of assimilating all-sky ABI radiance in improving the analysis of simulated BTs and radar reflectivity. In this subsection, simulated ABI channel 10 BTs of deterministic forecasts starting from the final analyses of the three DA experiments are compared to observations and the reference forecasts (i.e., CTRL and DETER) in Fig. 7 up to 4 h. Compared to CTRL and DETER, three DA experiments are able to better forecast the cloud patterns and evolutions of storms in the first 2 h of forecasts, although there are still displacement and intensity errors. The simulated cloudy BTs of all the experiments except DETER are overall warmer than observation, primarily near cloud edges, likely indicating weaker convections in these regions. In CTRL, cloud cover decreases in the first 2 h instead of increasing as observed, and the coverage starts to increase after 2 h (Figs. 7b2,b3). DETER and CTRL show similar cloud patterns, with DETER producing much larger cloud coverages and colder cloud-top BTs. Experiment HMX performs better than CTRL but not as well as MHX and MHX_15min, and not even compared to DETER, as storms become weaker after 2 h. Both CTRL and HMX fail to forecast the deep clouds in the

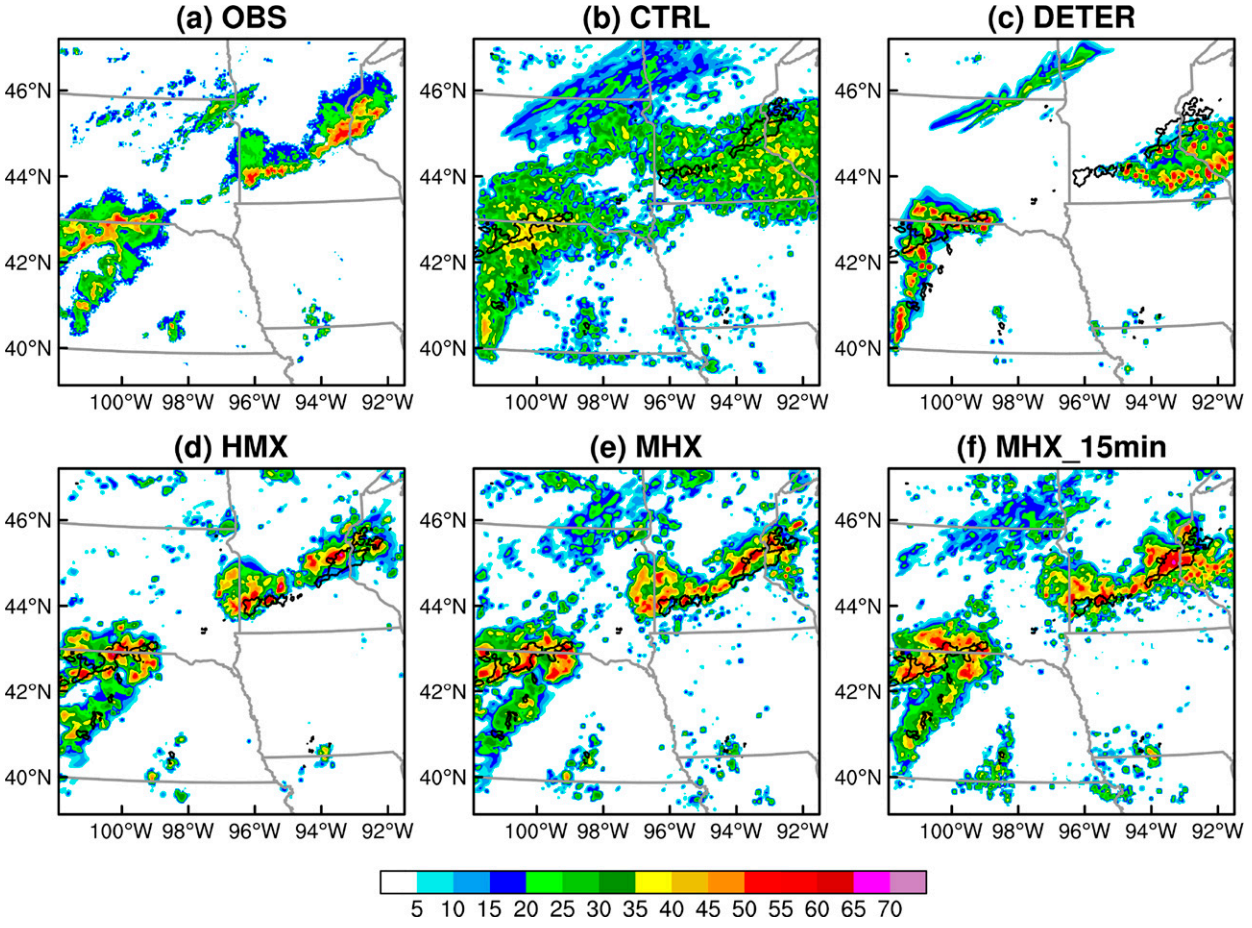

FIG. 6. (a) Observed and (b) simulated composite reflectivity (dBZ) for CTRL ensemble forecast mean, (c) DETER forecast, (d) ensemble analysis mean of experiment HMX, (e) MHX, and (f) MHX_15min valid at 2300 UTC, the time of final analysis. Black contours in (b)–(f) are for observed 40-dBZ composite reflectivity. Figure shows the entire WRF domain.

northeastern part of the convective line in the 3–4-h forecasts (Figs. 7b3,b4,d3,d4), while they are significantly overpredicted in DETER (Figs. 7c3,c4). MHX and MHX_15min are able to grow the deep clouds over the 4-h period, as in observations, and the clouds in eastern South Dakota cover slightly larger area (more consistent with observations) in MHX than in MHX_15min. The simulated clear-sky BTs are much higher in HMX at 1 h (Fig. 7d1) than in MHX (Fig. 7e1) and MHX_15min (Fig. 7f1) on the south side of the storms near the northeast end of the line, suggesting that the storm inflow air in that region is too dry and/or too warm. An inflow environment that is too dry, which is more likely due to the higher sensitivity of channel 10 to moisture than to temperature, will cause the forecasted storms to be weaker or to decay with time.

Figure 8 shows the time series of biases and RMSIs of ABI channels 10 and 14 for the forecasts calculated every 5 min. In terms of channel 10, the greatest improvements of assimilating all-sky ABI BTs are found with the forecast biases and RMSIs in the first hour, during which all three DA experiments show similar performance. The biases and RMSIs of the three DA experiment started to diverge somewhat after 0000 UTC (Fig. 8a). HMX is the closest to CTRL with larger negative biases and RMSIs than the other two DA experiments, and it underperforms DETER as well after ~0040 UTC. Both MHX and MHX_15min have smaller bias and RMSI than CTRL during the 4 h of forecast, but only during the first 2 h of forecast compared to DETER. Also, MHX_15min performs slightly better than MHX. For ABI window channel 14, the forecast errors of the experiments, especially those of CTRL and DETER, grow quickly during the first hour to magnitudes much larger than those of channel 10 due to the fact that channel 14 is more sensitive to cloud and surface parameters (Fig. 8b). In terms of channel 14 BT, CTRL and DETER exhibit more similar error characteristics and the differences between these two forecasts and the DA experiments are more notable. Unlike channel 10, the channel-14 bias and RMSI of MHX are the smallest among all experiments during almost the entire forecast period.

The 1–4-h composite reflectivity forecasts for all experiments are also compared with observations from the WSR-88D radars mapped onto model grid points. As is displayed in Fig. 9, the southwest–northeast precipitation band of the MCS can be reproduced by the DA experiments but is

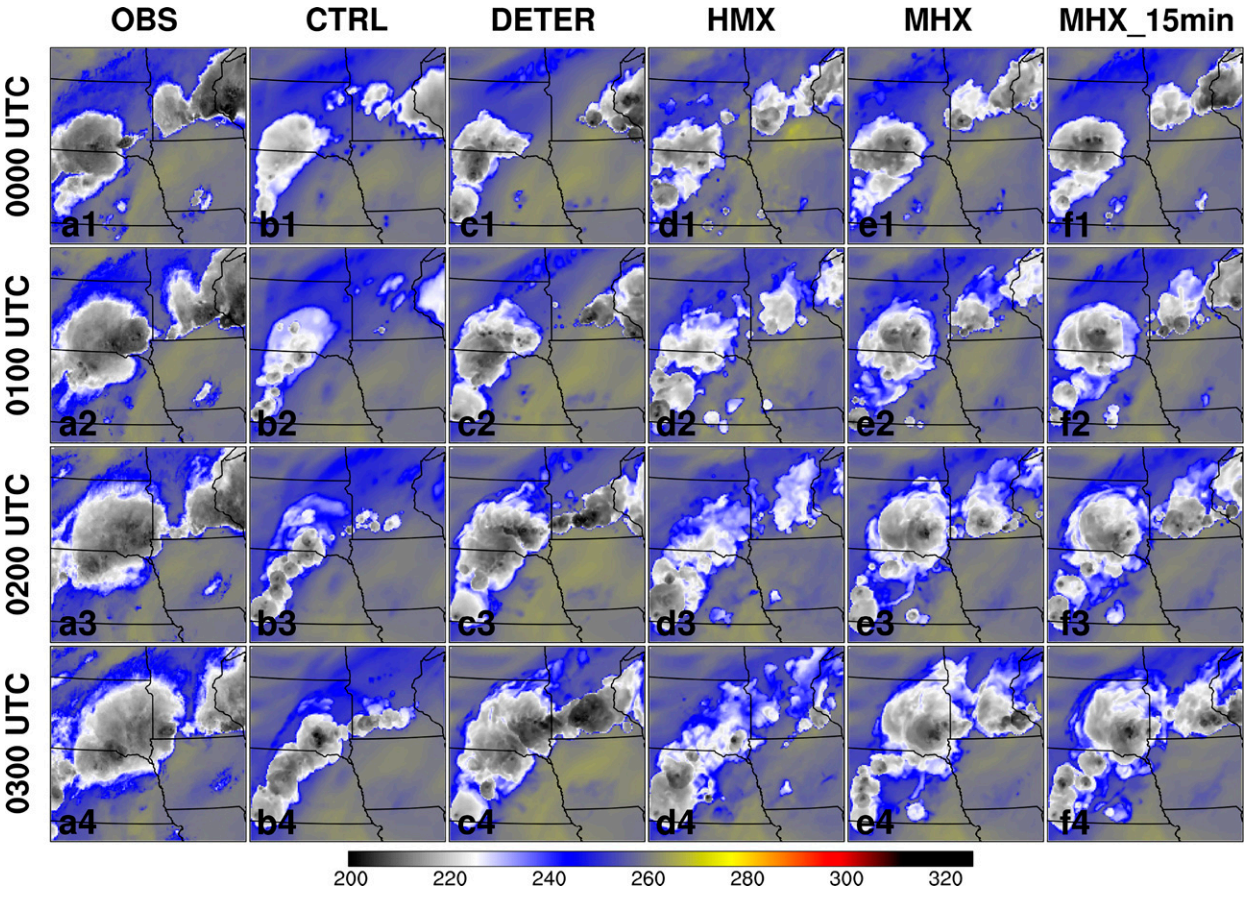

FIG. 7. (a1)–(a4) Observations and simulated brightness temperatures (K) of ABI channel 10 for (b1)–(b4) CTRL, (c1)–(c4) DETER, (d1)–(d4) HMX, (e1)–(e4) MHX, and (f1)–(f4) MHX_15min experiments after deterministic forecasts of 1 h in (a1)–(f1), 2 h in (a2)–(f2), 3 h in (a3)–(f3), and 4 h in (a4)–(f4) following 2-h DA (valid at 0000–0300 UTC). Figure shows the entire WRF domain.

missing in CTRL in the first 2 h of the forecast. In DETER, the displaced convective line in southern Minnesota in the final analysis (cf. Fig. 6c) decays while a new line develops slightly north of the observed one (Fig. 9c1). In the subsequent forecasts, the convective cores are stronger (Figs. 9c3,c4) than in DA experiments and observations. Compared to the analysis at the last EnKF cycle (Fig. 6) that has more intensive reflectivity with broader coverage than observed, the coverage

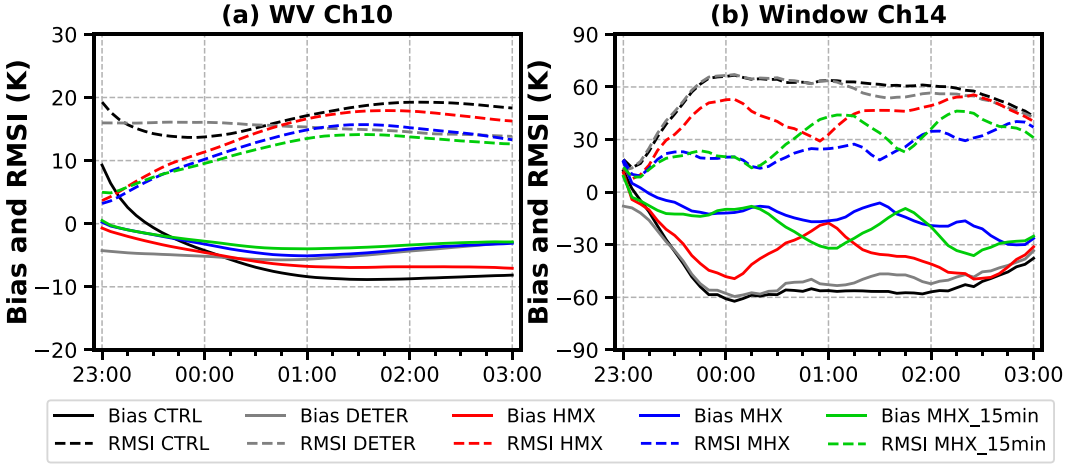

FIG. 8. Time series of biases $[\mathbf{y} - H(\overline{\mathbf{x}})]$ (solid curves) and RMSIs (dashed curves) of ABI (a) channel 10 and (b) channel 14 for the deterministic forecasts of experiments CTRL, DETER, HMX, MHX, and MHX_15min.

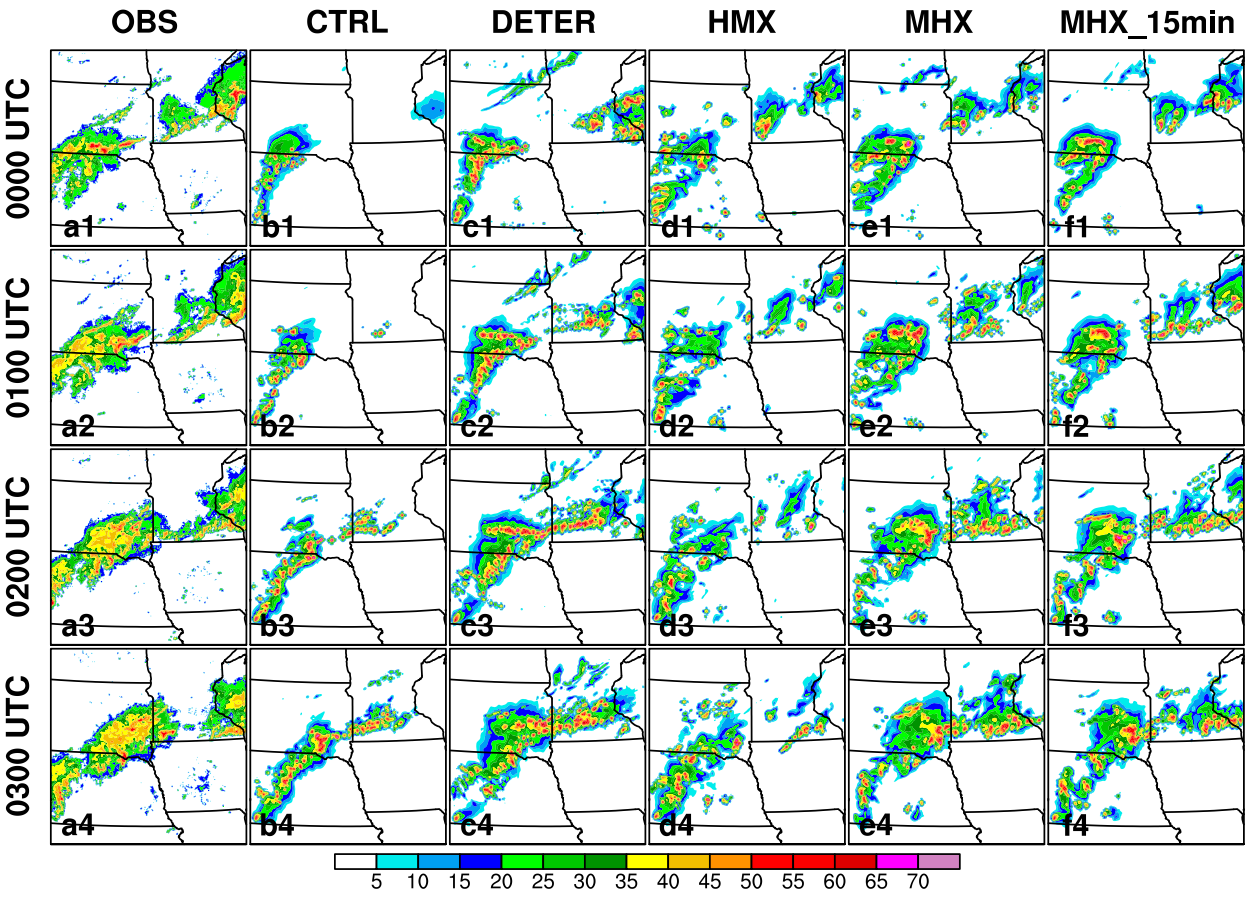

FIG. 9. As in Fig. 7, but for composite reflectivity fields from (a1)–(a4) observations, (b1)–(b4) CTRL, (c1)–(c4) DETER, (d1)–(d4) HMX, (e1)–(e4) MHX, and (f1)–(f4) MHX_15min.

of moderate-to-large reflectivity (i.e., >35 dBZ) in the forecast is generally smaller than in observations for all three DA experiments despite the still much higher reflectivity. In addition, the leading edge of the MCS in the northeast part is poorly maintained in the forecast of HMX while in MHX and MHX_15min the convective cores at the leading edge are much closer to observations through the 4 h of the forecast (e.g., Figs. 9e4,f4) and the forecast quality of MHX and MHX_15min are similar.

Neighborhood ETSs of the deterministic forecast for each experiment are computed with a 15-km neighborhood radius and are presented in Fig. 10. Generally, all three experiments that assimilate ABI BTs outperform CTRL notably for both

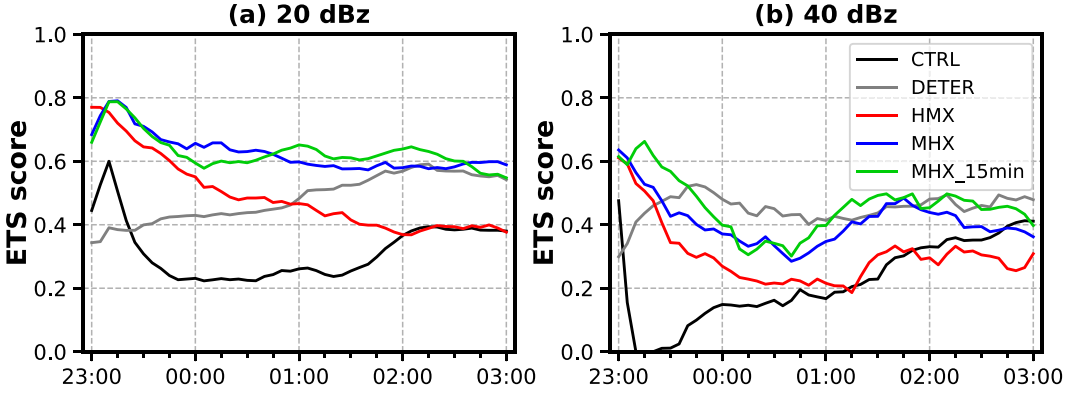

FIG. 10. Neighborhood ETSs of composite reflectivity forecasts calculated with a 15-km neighborhood radius at (a) 20- and (b) 40-dBZ thresholds for experiments CTRL (black), DETER (gray), HMX (red), MHX (blue), and MHX_15min (green).

weak and convective precipitation (in terms of weak and strong reflectivity), indicating improvements in both precipitation coverage and intensity forecast. For the threshold of 20 dBZ, CTRL experiences an increase in the first 10 min; this is because of decrease in the forecast of too large spatial coverage of weak reflectivity found in the initial condition that is the ensemble mean of 5-h forecasts. Such spindown of spurious weak echoes helps with the ETS in CTRL for a short time. After 10 min, the ETS of CTRL drops quickly in the first hour and then rises somewhat after ~0130 UTC when new convection develops (cf. Fig. 9). The improvement over DETER is, however, less obvious. ETS of DETER continuously increases during the 4 h of the forecast from a relatively low value of 0.35 at 2300 UTC to close to 0.6 at 0300 UTC.

For MHX and MHX_15min, there is also increase in ETS in the first 10 min (Fig. 10a) for the 20-dBZ threshold. There is also overestimation of weak reflectivity in the final analyses of these two experiments (Fig. 6e), and most of the weak reflectivity is cleared out during the first hour of forecast (Figs. 9d,e), improving the ETSs. Similar overestimation is not seen in the final analysis of HMX so an initial increase in ETS does not occur with HMX (Fig. 10a). The ETS of HMX decreases nearly monotonically with forecast time from its initial higher value than in other experiments, and even falls below that of DETER at about 0100 UTC and becomes essentially the same as that of CTRL after 0200 UTC. The drop in prediction skill is clearly related to the significant underprediction of storms in the northeast part of the MCS line (Figs. 9d2–d4). For MHX and MHX_15min, after the first 10 min, the ETSs decrease at a rate much slower than HMX and both are maintained at around 0.6 after 1 h. In terms of the 40-dBZ threshold corresponding to strong convection, CTRL has much lower scores in the first hour, after which it approaches that of HMX and exceeds its score somewhat after 0145 UTC (Fig. 10b). HMX underperforms MHX and MHX_15min throughout the entire forecast period. MHX_15min scores slightly higher than MHX because it somewhat better captures the convection in the northeast part of the MCS line (Fig. 9). It is worth noting that DETER shows comparable (or even higher) ETSs than MHX and MHX_15min after ~2330 UTC, probably because DETER overpredicts convective cores which benefits ETS scores. Overall, MHX and MHX_15min give consistently higher ETSs than HMX and CTRL for both 20- and 40-dBZ thresholds, and higher than DETER for 20 dBZ. In the next section, we will provide some explanations on the different behaviors of MHX and HMX.

d. *Error correlations and analysis increments in different experiments*

In this section, we try to gain some understanding on how the assimilation of BT data affects different state variables. Because the impact of the assimilation is largest in the first assimilation cycles (cf. Fig. 5), we examine the background error correlations between simulated BTs and the model state variables as calculated from the ensemble background at 2100 UTC, before the first EnKF analysis. Two pixels under different scene types are selected to illustrate the correlations, including a clear-sky pixel (marked with a star in Fig. 4b) and a cloudy pixel where background ensemble has spurious clouds (marked with a circle in Fig. 4b).

Figure 11 gives the vertical cross sections of correlations that pass through the representative locations of these two pixels. In clear-sky condition, the correlations between BT and model thermodynamic and dynamic variables show smooth patterns in the horizontal which are linked to the synoptic scale flows and conditions. The positive correlations between BT and potential temperatures mainly occur in the middle troposphere where the weighting function of ABI channel 10 peaks. The dominant negative correlations between BT and water vapor are related to the absorption of the outgoing longwave radiation. For horizontal wind components, the correlations are mostly associated with the horizontal transport of moisture. For example, in this case, the wind that comes from southeast in the low troposphere brings in drier air hence increasing the BT. Situations are quite different when it comes to the cloudy condition; the correlations become overall noisier mostly because of large variations of hydrometeors in space and across ensemble members. The upper boundary of positive (negative) correlations between BT and potential temperature (water vapor) occurs near the cloud top (marked by a circle). This is because infrared radiation is unable to penetrate clouds; therefore, it is most sensitive to cloud top temperature and water vapor when clouds are present. Asymmetrical features can be seen in dynamic fields resulting from the strong low-level inflow that comes from the southeast of the storm, helping enhance updraft hence having positive impact on convection strength. All hydrometeors show negative correlations with BT due to their absorption and scattering of radiation. The correlations between BT and ice particles (q_s, q_i, q_g) are generally stronger and broader in vertical than liquid clouds (q_r, q_c), indicating BT is more sensitive to ice particles. The spatial correlations between BT and model state variables in clear sky and cloudy regions shown suggest they are physical and such information is used by EnKF to update individual state variables using observed BT.

To examine the impact of actually assimilating ABI radiance on model state variables, cross sections of analysis increment for potential temperature and water vapor mixing ratio from HMX and MHX after the first EnKF cycle are compared in Fig. 12. The plots are in the same vertical cross sections as in Fig. 11. It is seen that the adjustments to state variables from clear sky BT observations are small for both HMX and MHX due to the small clear sky BT innovations (observation minus background simulation). Potential temperature analysis increments are within a range from −1 to 1 K, and water vapor adjustments are mainly under 500 hPa where moisture content is much higher (Figs. 12a,b). However, when the background ensemble mean is covered with spurious clouds, the lower background BTs compared to observations result in large positive prior innovations. Then based on the positive innovations and the positive (negative) background error correlations displayed in Fig. 11, model state variables show negative (positive) analysis increment accordingly, as shown in Figs. 12c and 12d. The increments are

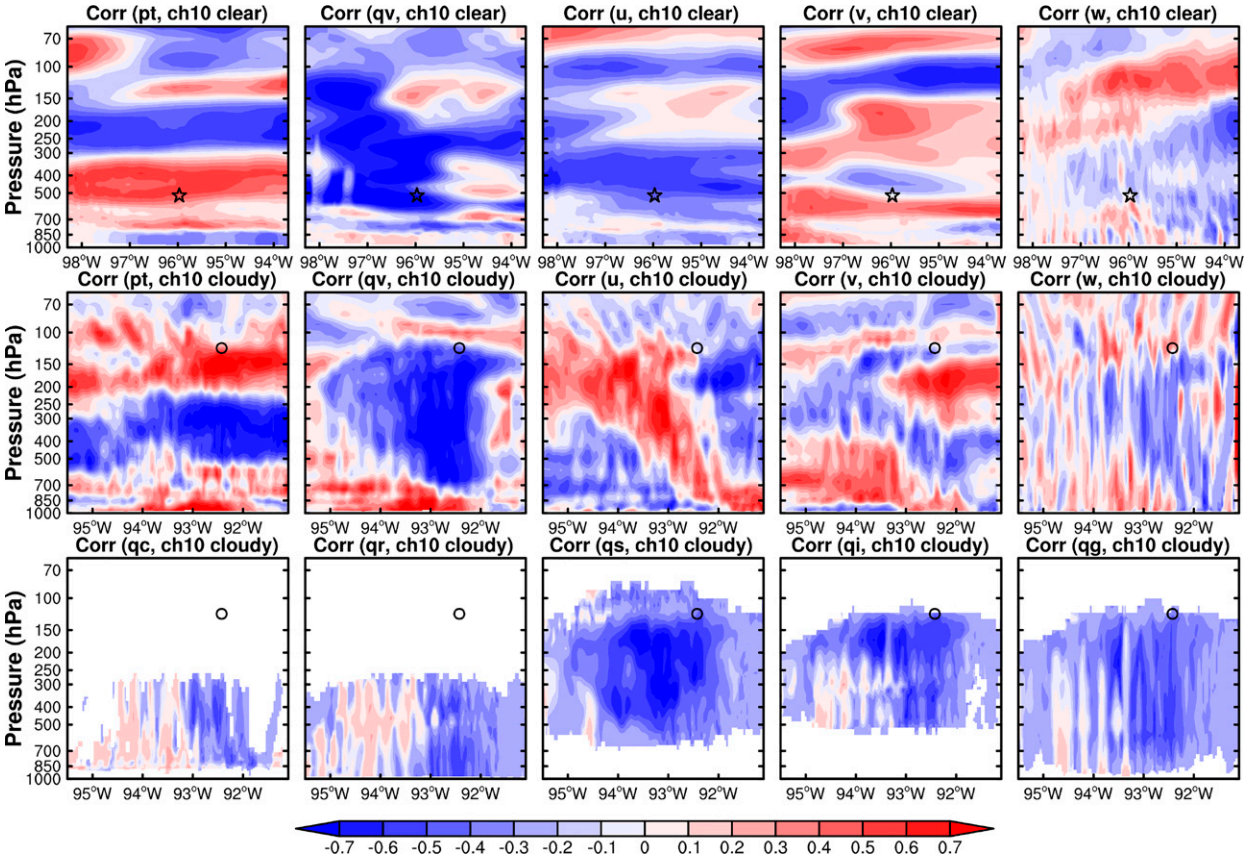

FIG. 11. East–west vertical cross sections of background error correlations between a simulated channel-10 brightness temperature under (top) clear-sky condition and (middle),(bottom) cloudy-sky condition, with model state variables potential temperature (p_t), water vapor mixing ratio (q_v), and wind components (u, v, w) in the top and middle rows and hydrometeors in the bottom row. The correlations are calculated from 3-h ensemble forecasts at 2100 UTC, the start time of DA cycles. The clear and cloudy brightness temperature locations are marked as a star and circle, respectively, and they are placed at the level of their respective peak response function.

noticeably larger in cloudy sky than in clear sky. Also, under cloudy sky both potential temperature and water vapor increments in HMX are generally larger in magnitude than those in MHX. In terms of adjustment to hydrometeors, both HMX and MHX work well in clearing spurious clouds as most clouds are removed after the first EnKF cycle (see Fig. 13). Clearly, HMX outperforms MHX in removing both frozen and liquid hydrometeors because of its larger innovations calculated based on the specific $H(\mathbf{x})$ formula. It should also be pointed out that assimilating all-sky ABI radiance clearly has more impact on the analysis of frozen hydrometeors than liquid hydrometeors, because of the higher sensitivity of ABI radiance to upper-level hydrometeors based on the radiative transfer model.

It is evident that in HMX, assimilating ABI radiance may produce larger analysis increments than in MHX when the background is covered with spurious clouds. However, the effect of these large increments still needs to be discussed since erroneous analysis increments could result in model imbalances. Figure 14 shows the 850-hPa water vapor mixing ratio from CTRL and the difference between the water vapor of each DA experiment and CTRL at 2100, 2200, and 2300 UTC. After the first EnKF cycle, some water vapor in the northeastern domain is reduced in HMX together with the removal of spurious model clouds in this area, but this reduction is not seen in MHX or MHX_15min. Small positive differences (or analysis increments) are observed for all three DA experiments over southern Minnesota. They are probably related to the less reliable background error covariance related to limited ensemble size and errors in the mean state. In the subsequent DA cycles, those positive differences become negative in HMX and MHX, and are moderate in MHX_15min, further suppressing the development of spurious convection. Meanwhile, the atmosphere becomes more moist over the observed cloudy region (e.g., in western Nebraska), contributing to redevelopment of new storms.

The difference between the DA experiments is also noticeable. For example, in HMX, large negative differences become dominant at the final analysis time (Fig. 14j) especially over the forward region of the leading convective edge in southern Minnesota. The environment field in MHX is overall more moist than in HMX, but negative differences still occur in the northeastern domain. In MHX_15min, smaller adjustments are made to the same region, however. This significant

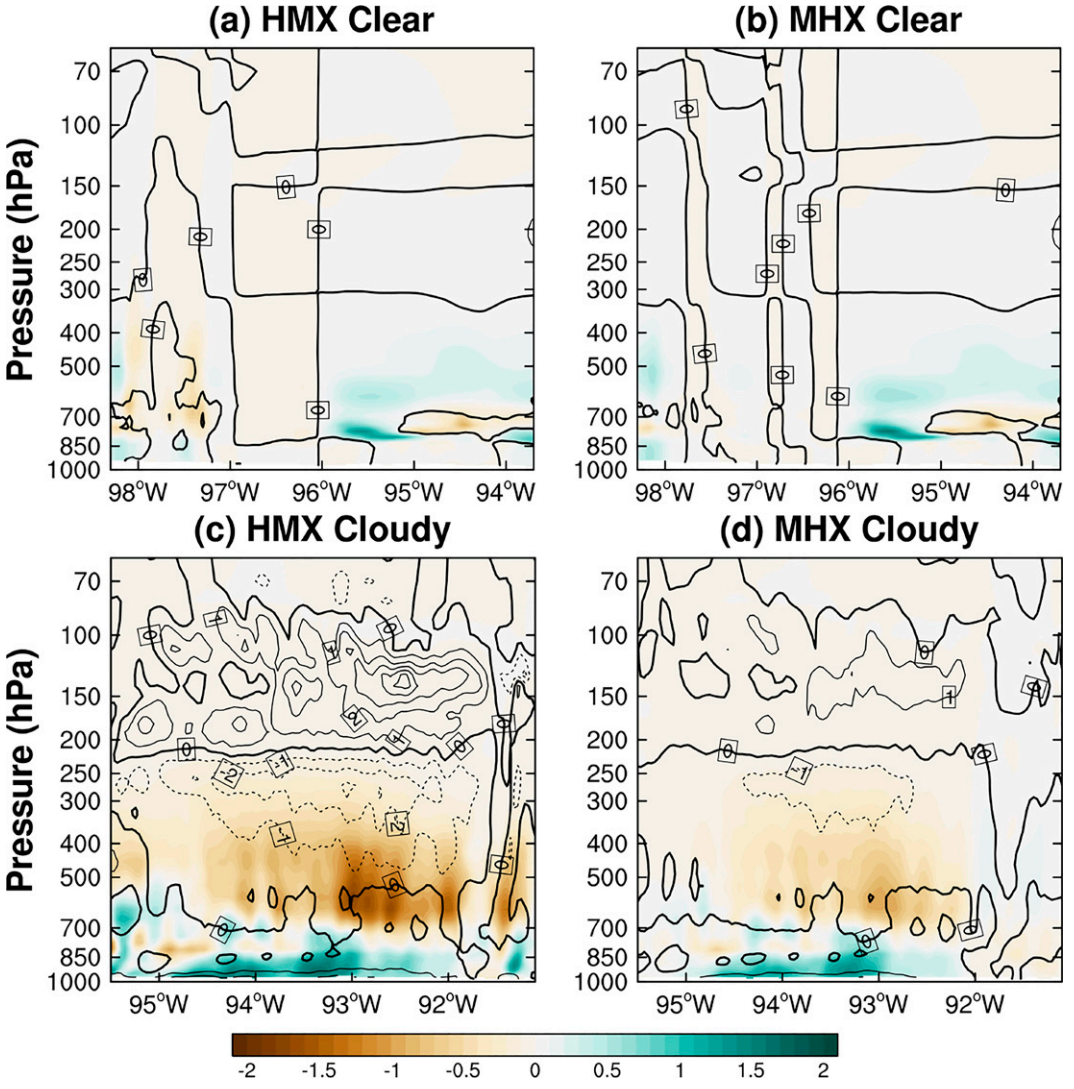

FIG. 12. Vertical cross sections of analysis increments of water vapor mixing ratio (g kg^{-1}; shading) and potential temperature (contours; 1-K interval) for experiments (left) HMX and (right) MHX at 2100 UTC, the time of the first analysis. The cross sections pass through the same clear- and cloudy-sky observation locations in Fig. 11.

overdrying behavior is probably responsible for the underperformance of HMX. On the other hand, a 15-min DA interval provides more time for adjustment in the model, possibly reducing imbalances among state variables and leading to a comparable forecast performance as the 5-min interval experiment.

Figure 15 shows the time series of the domain-averaged 500- and 850-hPa forecast and analysis ensemble mean water vaper mixing ratio (g kg^{-1}) for each experiment during throughout the EnKF cycles. It is seen that MHX and MHX_15min have higher domain-averaged water vapor at 500 hPa than CTRL and HMX. This is due to more moisture removal in HMX than in MHX and MHX_15min at this level following the clearing of spurious model clouds in the first few of EnKF cycles, while assimilating observations in cloud-free area moistens the 500 hPa environment. After the first few cycles, EnKF generally decreases the 500-hPa moisture while ensemble forecasts increase the moisture slightly (by transporting low-level moisture to the level through convection). Chan et al. (2020) also saw that the moistening effects in forecast steps are unable to counter the drying effects brought by ABI assimilation, leading to the overall drying. The drying is more in MHX than in MHX_15min (Fig. 15a) due to 3 times more DA cycles. In HMX, this is a small reduction instead of large increase in the first analysis, leading to an overall much lower moisture level than in the other two experiments.

At 850 hPa, all three DA experiments show lower domain-averaged water vapor than CTRL and the differences become more significant with time. The largest decreases in water vapor result from the EnKF updates in the first two cycles; these are mainly associated with the clearing of spurious clouds in the background, especially those in southern Minnesota (see

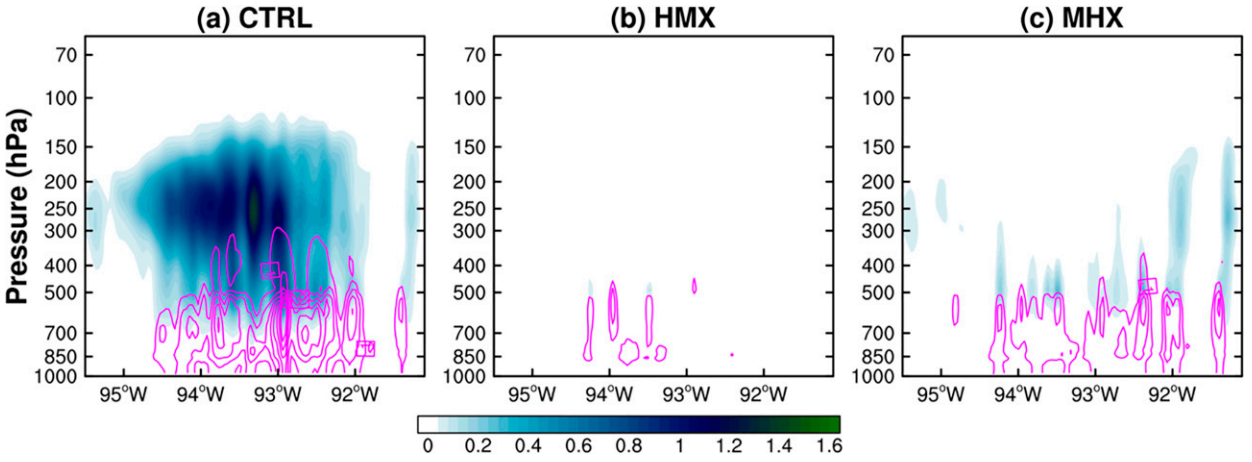

FIG. 13. Vertical cross sections of the mixing ratio (g kg^{-1}) of frozen hydrometeors ($q_s + q_i + q_g$; shading) and liquid hydrometeors ($q_c + q_r$; contours) for (a) CTRL ensemble mean forecast, (b) HMX mean analysis, and (c) MHX mean analysis valid at 2100 UTC. The cross section passes through the same cloudy-sky observation location in Fig. 11.

Fig. 4). Unlike at 500 hPa, drying also occurs in the forecast steps at 850 hPa, and is the strongest in HMX. Investigation shows that the clearing of clouds by EnKF is accompanied by creation and enhancement of low-level downward motion (not shown), and the enhancement is larger in HMX. The downward advection of moisture in such regions leads to the drying during the forecast steps. MHX_15min is overall moister at both low- and midlevels due to a smaller number of EnKF cycles.

To better understand the large difference in the behaviors of $H(\overline{\mathbf{x}})$ and $\overline{H(\mathbf{x})}$ formulations in the ensemble mean state update equation of EnSRF, the BTs calculated according to $H(\overline{\mathbf{x}})$ and $\overline{H(\mathbf{x})}$ from the background ensemble forecasts before the first DA cycle at 2100 UTC are presented in Fig. 16, along with the ensemble mean and standard deviation of total hydrometeor mixing ratio at the 200-hPa level, which is close to the cloud top. It is seen that in the clear sky condition, the differences of $H(\overline{\mathbf{x}})$ and $\overline{H(\mathbf{x})}$ are relatively small, due to the weak nonlinearity of the observation operator in these regions, leading to the small differences in analysis increments between HMX and MHX, as seen earlier in Figs. 12a and 12b. When clouds are present, however, $H(\overline{\mathbf{x}})$ are much lower than $\overline{H(\mathbf{x})}$. This is because there is a large spread in the cloudiness conditions among the ensemble members, as indicated by the near-cloud-top total hydrometeor mixing ratio spread (which is as large as the mean value). The calculated BTs from $H(\mathbf{x})$ also have a large spread; therefore, their mean values are not that low (Fig. 16b). When the mean BT is calculated from $H(\overline{\mathbf{x}})$, even though the mean of total hydrometeors $\overline{\mathbf{x}}$ may be significantly reduced by the ensemble averaging, the calculated BT can still be low, as long upper-level clouds are present. As a result, large positive BT innovations are found where background ensemble contains spurious convection, even if there is significant spread among the ensemble members. This results in larger analysis increments in HMX than in MHX (see Figs. 12c,d and 13). Due to the large diversity of clouds among ensemble members after 3 h of spinup ensemble forecasts, the differences between $H(\overline{\mathbf{x}})$ and $\overline{H(\mathbf{x})}$ in cloudy regions are more noticeable in the first few cycles before most spurious clouds generated in the spinup period are cleared. The removal of water vapor through cross-covariance in HMX is also larger than in MHX, leading to the too much low-level drying in HMX noticed above.

5. Summary and conclusions

In this study, all-sky IR BTs at their native resolution from the GOES-R ABI water vapor channel 10 are assimilated into the 3-km WRF model using an enhanced version of GSI EnKF, and the assimilation impacts on the analysis and forecast of a mesoscale convective system (MCS) are examined. Adding to a limited number of prior studies assimilating high-resolution GOES-R ABI BT type of data, this study employs the operational GSI EnKF framework and tries to optimize certain DA configurations and gain an understanding of the behaviors of BT DA in cloudy and clear-sky regions.

A radius of 30 km was found to perform well when both clear and cloudy ABI BTs are assimilated, yielding smaller analysis biases and RMSIs as well as better reflectivity forecast than other radii examined. This "optimal" choice of localization radius is in a general agreement with those used in most recent studies (e.g., Cintineo et al. 2016; Zhang et al. 2018; Otkin and Potthast 2019; Jones et al. 2020).

Because of the high nonlinearity of the radiative transfer model that is used as the BT observation operator, especially in cloudy regions, the DA results are found to be very sensitive to the way the ensemble mean observation prior is calculated. Two different formulations of the ensemble mean observation prior in the ensemble mean update equation are examined. Using the 30-km localization radius, three DA experiments are carried out. HMX uses the original way of calculating the ensemble mean observation prior in GSI EnKF, which applies the observation operator to the ensemble mean state, i.e., using $H(\overline{\mathbf{x}})$, in the ensemble mean state update

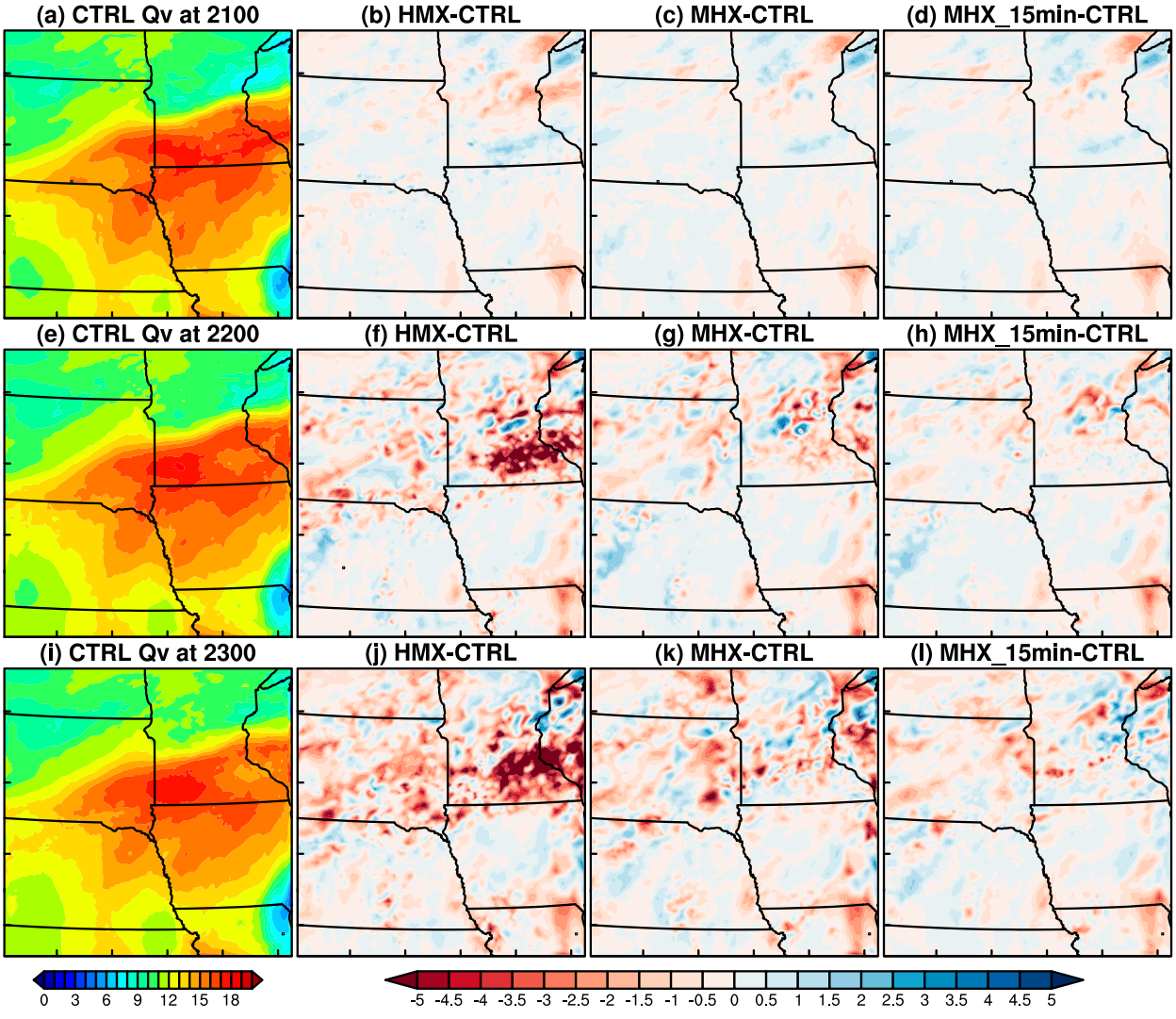

FIG. 14. The 850-hPa water vapor mixing ratio (g kg^{-1}) from experiment CTRL at (a) 2100, (e) 2200, and (i) 2300 UTC. The 850-hPa water vapor mixing ratio (g kg^{-1}) differences between CTRL and experiments (b),(f),(j) HMX; (c),(g),(k) MHX; and (d),(h),(l) MHX_15min at 2100 UTC in (b)–(d), 2200 UTC in (f)–(h), and 2300 UTC in (j)–(l).

equation. MHX replaces $H(\overline{\mathbf{x}})$ with $\overline{H(\mathbf{x})}$ that performs the mean operation on the ensemble of observation priors. MHX_15min is the same as MHX except that BTs are assimilated every 15 min instead of 5 min.

The results show that assimilating ABI BTs can, despite some problems, build up observed storms within the model and remove spurious storms in the background, as was also previously found in Zhang et al. (2018). Both biases and RMSIs of the background and analysis BTs are significantly reduced during the assimilation cycles with MHX generally has the smallest biases and RMSIs. Ensemble spread during the DA cycles is well maintained by applying the relaxation to prior spread covariance inflation method, and MHX produces a closer-to-optimal prior consistency ratio. The assimilation of ABI BTs also improves simulated radar reflectivity in the final analyses, despite more spurious precipitation is analyzed in MHX_15min.

Simulated channel 10 BTs of the 4-h deterministic forecasts show that assimilating all-sky ABI BTs improves the forecast of simulated BTs. Forecast biases and RMSIs of both water vapor channel-10 and window channel-14 BTs are reduced in the three DA experiments to some extent compared to CTRL and DETER that do not assimilate any data. MHX has the smallest forecast errors, slightly outperforming MHX_15min, while HMX performs the worst. The forecast of weak radar reflectivity is also significantly improved compared to CTRL and DETER, up to 4 and 2 h, respectively. Convection is more intensive and organized, closer to observations, in MHX and MHX_15min than in HMX, hence have higher reflectivity ETSs. It is also noted that increasing assimilating frequency from 15 to 5 min, at least in this case using a 2-h assimilation window, does not significantly improve the forecasting results. This result is unlike the finding in Zhang et al. (2021) that increasing assimilating frequency of ABI observations from

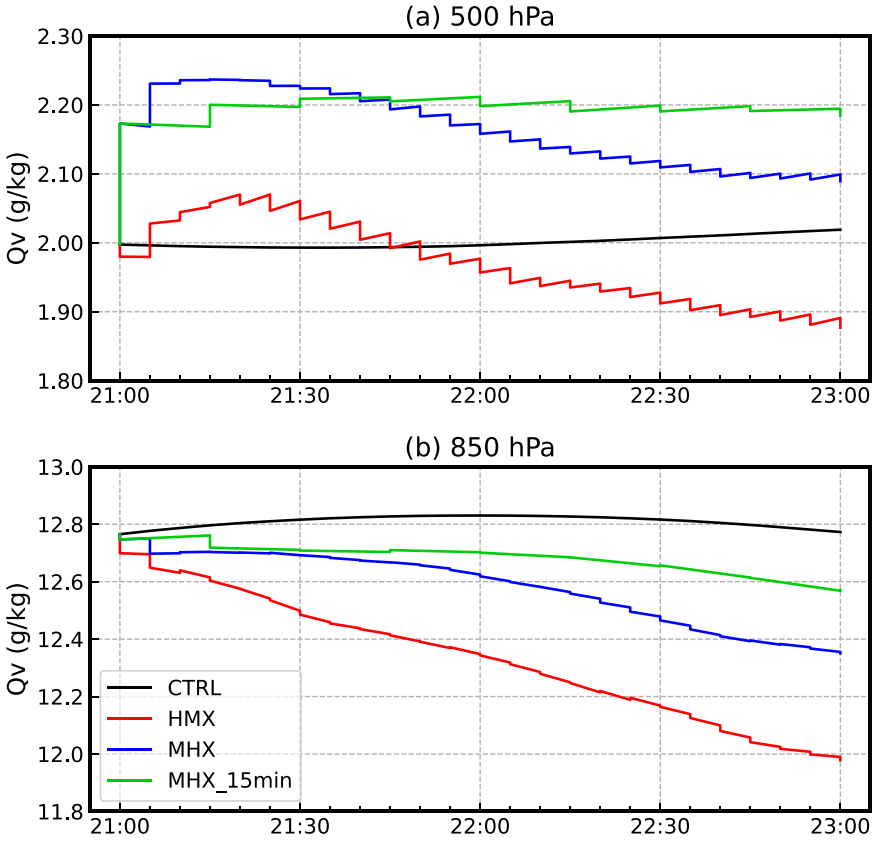

FIG. 15. Time series of domain-averaged (a) 500- and (b) 850-hPa water vaper mixing ratio (g kg^{-1}) for experiments CTRL (black), HMX (red), MHX (blue), and MHX_15min (green) during the entire EnKF cycles.

10 to 5 min improves both analysis and forecast of a thunderstorm, thus further investigations are needed.

Analyses of background error correlations before the first EnKF cycle indicate strong correlations between simulated BT and model state variables in clear and cloudy conditions. The correlations between BT and thermodynamic and dynamic fields in clear sky conditions are smooth in space and are mainly associated with the synoptic-scale flows, while those in cloudy conditions have much more small-scale structures. The adjustments of model state variables are consistent with the sign and magnitude of background error covariance and prior innovations. Larger observation innovations in HMX lead to larger analysis increment of hydrometeors, which is beneficial in terms of clearing spurious model clouds. However, the clearing of spurious clouds is also accompanied by the reduction in water vapor, mostly through background error spatial covariance. Larger positive innovations in HMX may result in excessive corrections to thermodynamic and dynamic fields, causing incorrect storm decay in the forecast; this is so despite of the use of an adaptive observation error inflation procedure that significantly increase the assumed observation error used in EnKF when observation innovation is large. In experiments MHX and MHX_15min that use $\overline{H(\mathbf{x})}$ instead of $H(\overline{\mathbf{x}})$ in the calculation of observation prior mean in EnKF, the moisture drying is reduced.

Although the results presented in this study show positive impacts of assimilating all-sky ABI observations, there remain issues and limitations with the study. First, over drying of the low levels when removing spurious clouds/convection in the forecast background is an important issue and such overdrying needs to be reduced. Apart from the effect on the calculation of BT observation prior of the ensemble mean, the nonlinearity of the CRTM observation operator also leads to differences between the "scalable" implementation of EnSRF used in this study (and in GSI EnKF) and the original serial EnSRF algorithm. The former updates observation prior $\mathbf{y}^b$ using the "scalable" implementation of (Anderson and Collins 2007) according to Eq. (3) while the latter recalculating $\mathbf{y}^b$ using CRTM from state $\mathbf{x}$ updated by earlier observations. This difference can also affect the observation innovation ($\mathbf{y}^o - \mathbf{y}^b$) and affect the updating of moisture fields.

The assimilated ABI observations used in this study are at their native resolutions. Further sensitivity tests on data thinning intervals should be performed to see if computational costs can be reduced without much loss of analysis quality and/or help alleviate the effect of possible spatial error correlations among observations. Also, more optimization on horizontal and vertical localization radii, possibly using different radii in cloudy and clear regions, can be performed to take into account the different characteristics of clear and cloud

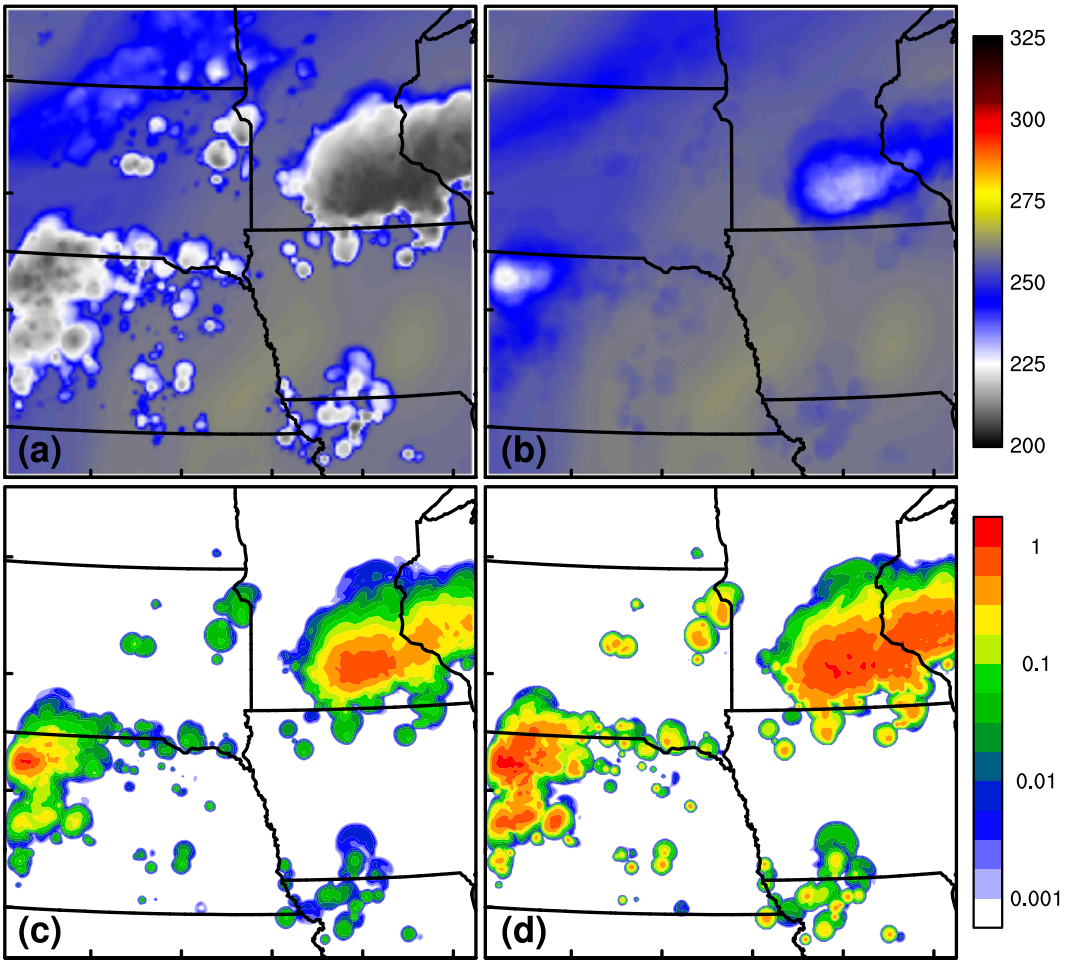

FIG. 16. Brightness temperature calculated according to (a) $H(\overline{\mathbf{x}})$ and (b) $\overline{H(\mathbf{x})}$ from the background ensemble forecasts at 2100 UTC, the start time of DA cycles, and the (c) ensemble mean and (d) standard deviation of total hydrometeor mixing ratio at the 200-hPa level, which is near the height of cloud tops in southeastern Minnesota.

sky observations. The ensemble is generally underdispersive during the DA cycles, and the relaxation-to-prior-spread inflation method may need to be supplemented by other inflation methods. Moreover, model and observation biases are not considered in this study. Applying bias correction to ABI BTs has the potential to improve our results although bias correction in the cloudy regions remains a very challenging task because the errors in the simulated background BTs can be very large and uncertain. The adaptive observation error inflation method employed, while very helpful, is ad hoc at the best, and it acts to reduce the observation impact when observation innovation is large. Better treatments and at least tuning of the method may be needed to further improve results. Further, assimilating multichannel ABI observations with additional quality control, and the assimilation of ABI radiances together with other observations such as high-resolution surface observations, radar observations and even GOES lighting data could provide more accurate analyses of storm structures and their environment. Finally, the results of in this paper are based on a single case; the generalization of the some of the conclusions will require testing with a large enough sample. These are some of the issues needing further research before all-sky GOES-R ABI type observations can be effectively assimilated into operational convective-scale forecasting systems with robust positive impact.

Acknowledgments. This work was primarily sponsored by the National Key Research and Development Program of China under Grant 2017YFC1502102, and the National Natural Science Foundation of China under Grant 41805071. The first author acknowledges support from the Chinese Scholarship Council for the enabled visit to the University of Oklahoma where most of this research was conducted. The NSF Extreme Science and Engineering Discovery Environment (XSEDE) supercomputing resources at the Texas Advanced Computing Center were used for most of the DA experiments.

Data availability statement. The GOES-R data used in this manuscript can be downloaded from NOAA's Comprehensive Large Array-Data Stewardship System (CLASS) at

https://www.class.noaa.gov. The final version of data used and presented in this study are available at https://doi.org/10.7910/DVN/PDYLFS.

REFERENCES

Aksoy, A., D. C. Dowell, and C. Snyder, 2009: A multicase comparative assessment of the ensemble Kalman filter for assimilation of radar observations. Part I: Storm-scale analyses. *Mon. Wea. Rev.*, **137**, 1805–1824, https://doi.org/10.1175/2008MWR2691.1.

Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, **129**, 2884–2903, https://doi.org/10.1175/1520-0493(2001)129<2884:AEAKFF>2.0.CO;2.

——, and N. Collins, 2007: Scalable implementations of ensemble filter algorithms for data assimilation. *J. Atmos. Oceanic Technol.*, **24**, 1452–1463, https://doi.org/10.1175/JTECH2049.1.

Bessho, K., and Coauthors, 2016: An introduction to Himawari-8/9—Japan's new-generation geostationary meteorological satellites. *J. Meteor. Soc. Japan*, **94**, 151–183, https://doi.org/10.2151/jmsj.2016-009.

Chan, M.-Y., F. Zhang, X. Chen, and L. R. Leung, 2020: Potential impacts of assimilating all-sky satellite infrared radiances on convection-permitting analysis and prediction of tropical convection. *Mon. Wea. Rev.*, **148**, 3203–3224, https://doi.org/10.1175/MWR-D-19-0343.1.

Chen, F., and J. Dudhia, 2001: Coupling an advanced land surface-hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Wea. Rev.*, **129**, 569–585, https://doi.org/10.1175/1520-0493(2001)129<0569:CAALSH>2.0.CO;2.

Chen, Y., H. Wang, J. Min, X.-Y. Huang, P. Minnis, R. Zhang, J. Haggerty, and R. Palikonda, 2015: Variational assimilation of cloud liquid/ice water path and its impact on NWP. *J. Appl. Meteor. Climatol.*, **54**, 1809–1825, https://doi.org/10.1175/JAMC-D-14-0243.1.

Cintineo, R., J. A. Otkin, M. Xue, and F. Kong, 2014: Evaluating the performance of planetary boundary layer and cloud microphysical parameterization schemes in convection-permitting ensemble forecasts using synthetic *GOES-13* satellite observations. *Mon. Wea. Rev.*, **142**, 163–182, https://doi.org/10.1175/MWR-D-13-00143.1.

——, ——, T. A. Jones, S. Koch, and D. J. Stensrud, 2016: Assimilation of synthetic GOES-R ABI infrared brightness temperatures and WSR-88D radar observations in a high-resolution OSSE. *Mon. Wea. Rev.*, **144**, 3159–3180, https://doi.org/10.1175/MWR-D-15-0366.1.

Derber, J. C., and W.-S. Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. *Mon. Wea. Rev.*, **126**, 2287–2299, https://doi.org/10.1175/1520-0493(1998)126<2287:TUOTCC>2.0.CO;2.

Dowell, D. C., and L. J. Wicker, 2009: Additive noise for storm-scale ensemble data assimilation. *J. Atmos. Oceanic Technol.*, **26**, 911–927, https://doi.org/10.1175/2008JTECHA1156.1.

——, F. Q. Zhang, L. J. Wicker, C. Snyder, and N. A. Crook, 2004: Wind and temperature retrievals in the 17 May 1981 Arcadia, Oklahoma, supercell: Ensemble Kalman filter experiments. *Mon. Wea. Rev.*, **132**, 1982–2005, https://doi.org/10.1175/1520-0493(2004)132<1982:WATRIT>2.0.CO;2.

——, L. J. Wicker, and C. Snyder, 2011: Ensemble Kalman filter assimilation of radar observations of the 8 May 2003 Oklahoma City supercell: Influences of reflectivity observations on storm-scale analyses. *Mon. Wea. Rev.*, **139**, 272–294, https://doi.org/10.1175/2010MWR3438.1.

Ebert, E. E., 2009: Neighborhood verification: A strategy for rewarding close forecasts. *Wea. Forecasting*, **24**, 1498–1510, https://doi.org/10.1175/2009WAF2222251.1.

Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J. Geophys. Res.*, **99**, 10 143–10 162, https://doi.org/10.1029/94JC00572.

——, 2009: *Data Assimilation: The Ensemble Kalman Filter*. 2nd ed. Springer, 307 pp.

Eyre, J. R., G. A. Kelly, A. P. McNally, E. Andersson, and A. Persson, 1993: Assimilation of TOVS radiance information through one-dimensional variational analysis. *Quart. J. Roy. Meteor. Soc.*, **119**, 1427–1463, https://doi.org/10.1002/qj.49711951411.

Fabry, F., and J. Z. Sun, 2010: For how long should what data be assimilated for the mesoscale forecasting of convection and why? Part I: On the propagation of initial condition errors and their implications for data assimilation. *Mon. Wea. Rev.*, **138**, 242–255, https://doi.org/10.1175/2009MWR2883.1.

Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, **125**, 723–757, https://doi.org/10.1002/qj.49712555417.

Geer, A. J., P. Bauer, and P. Lopez, 2010: Direct 4D-Var assimilation of all-sky radiances. Part II: Assessment. *Quart. J. Roy. Meteor. Soc.*, **136**, 1886–1905, https://doi.org/10.1002/qj.681.

——, and Coauthors, 2018: All-sky satellite data assimilation at operational weather forecasting centres. *Quart. J. Roy. Meteor. Soc.*, **144**, 1191–1217, https://doi.org/10.1002/qj.3202.

Han, Y., P. V. Delst, Q. Liu, F. Weng, B. Yan, R. Treadon, and J. Derber, 2006: JCSDA Community Radiative Transfer Model (CRTM): Version 1. NOAA Tech. Rep. NESDIS 122, 40 pp.

Honda, T., and Coauthors, 2018: Assimilating all-sky *Himawari-8* satellite infrared radiances: A case of Typhoon Soudelor (2015). *Mon. Wea. Rev.*, **146**, 213–229, https://doi.org/10.1175/MWR-D-16-0357.1.

Hong, S.-Y., Y. Noh, and J. Dudhia, 2006: A new vertical diffusion package with an explicit treatment of entrainment processes. *Mon. Wea. Rev.*, **134**, 2318–2341, https://doi.org/10.1175/MWR3199.1.

Houtekamer, P. L., and H. L. Mitchell, 2001: A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, **129**, 123–137, https://doi.org/10.1175/1520-0493(2001)129<0123:ASEKFF>2.0.CO;2.

Hunt, B. R., E. J. Kostelich, and I. Szunyogh, 2007: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D*, **230**, 112–126, https://doi.org/10.1016/j.physd.2006.11.008.

Iacono, M. J., J. S. Delamere, E. J. Mlawer, M. W. Shephard, S. A. Clough, and W. D. Collins, 2008: Radiative forcing by long-lived greenhouse gases: Calculations with the AER radiative transfer models. *J. Geophys. Res.*, **113**, D13103, https://doi.org/10.1029/2008JD009944.

Janjić, Z. I., 1994: The step-mountain eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes. *Mon. Wea. Rev.*, **122**, 927–945, https://doi.org/10.1175/1520-0493(1994)122<0927:TSMECM>2.0.CO;2.

——, 1996: The surface layer in the NCEP Eta model. *11th Conf. on Numerical Weather Prediction*, Norfolk, VA, Amer. Meteor. Soc., 354–355.

Jiménez, P. A., and J. Dudhia, 2012: Improving the representation of resolved and unresolved topographic effects on surface wind in the WRF Model. *J. Appl. Meteor. Climatol.*, **51**, 300–316, https://doi.org/10.1175/JAMC-D-11-084.1.

Jones, T. A., and D. J. Stensrud, 2015: Assimilating cloud water path as a function of model cloud microphysics in an idealized simulation. *Mon. Wea. Rev.*, **143**, 2052–2081, https://doi.org/10.1175/MWR-D-14-00266.1.

——, ——, P. Minnis, and R. Palikonda, 2013a: Evaluation of a forward operator to assimilate cloud water path into WRF-DART. *Mon. Wea. Rev.*, **141**, 2272–2289, https://doi.org/10.1175/MWR-D-12-00238.1.

——, J. A. Otkin, D. J. Stensrud, and K. Knopfmeier, 2013b: Assimilation of satellite infrared radiances and Doppler radar observations during a cool season observing system simulation experiment. *Mon. Wea. Rev.*, **141**, 3273–3299, https://doi.org/10.1175/MWR-D-12-00267.1.

——, ——, ——, and ——, 2014: Forecast evaluation of an observing system simulation experiment assimilating both radar and satellite data. *Mon. Wea. Rev.*, **142**, 107–124, https://doi.org/10.1175/MWR-D-13-00151.1.

——, X. Wang, P. Skinner, A. Johnson, and Y. Wang, 2018: Assimilation of *GOES-13* imager clear-sky water vapor (6.5 μm) radiances into a Warn-on-Forecast System. *Mon. Wea. Rev.*, **146**, 1077–1107, https://doi.org/10.1175/MWR-D-17-0280.1.

——, and Coauthors, 2020: Assimilation of *GOES-16* radiances and retrievals into the Warn-on-Forecast System. *Mon. Wea. Rev.*, **148**, 1829–1859, https://doi.org/10.1175/MWR-D-19-0379.1.

Kain, J. S., and Coauthors, 2010: Assessing advances in the assimilation of radar data and other mesoscale observations within a collaborative forecasting-research environment. *Wea. Forecasting*, **25**, 1510–1521, https://doi.org/10.1175/2010WAF2222405.1.

Kleist, D. T., D. F. Parrish, J. C. Derber, R. Treadon, W.-S. Wu, and S. Lord, 2009: Introduction of the GSI into the NCEP global data assimilation system. *Wea. Forecasting*, **24**, 1691–1705, https://doi.org/10.1175/2009WAF2222201.1.

Kong, R., M. Xue, and C. Liu, 2018: Development of a hybrid En3DVar data assimilation system and comparisons with 3DVar and EnKF for radar data assimilation with observing system simulation experiments. *Mon. Wea. Rev.*, **146**, 175–198, https://doi.org/10.1175/MWR-D-17-0164.1.

——, ——, A. O. Fierro, Y. Jung, C. Liu, E. R. Mansell, and D. R. MacGorman, 2020: Assimilation of *GOES-16* Geostationary Lightning Mapper flash extent density data in GSI EnKF for the analysis and short-term forecast of a mesoscale convective system. *Mon. Wea. Rev.*, **148**, 2111–2133, https://doi.org/10.1175/MWR-D-19-0192.1.

Köpken, C., G. Kelly, and J.-N. Thépaut, 2004: Assimilation of Meteosat radiance data within the 4D-Var system at ECMWF: Assimilation experiments and forecast impact. *Quart. J. Roy. Meteor. Soc.*, **130**, 2277–2292, https://doi.org/10.1256/qj.02.230.

Labriola, J., Y. Jung, C. Liu, and M. Xue, 2021: Evaluating forecast performance and sensitivity to the GSI EnKF data assimilation configuration for the 28–29 May 2017 mesoscale convective system case. *Wea. Forecasting*, **36**, 127–146, https://doi.org/10.1175/WAF-D-20-0071.1.

Ma, Z., E. S. Maddy, B. Zhang, T. Zhu, and S. A. Boukabara, 2017: Impact assessment of *Himawari-8* AHI data assimilation in NCEP GDAS/GFS with GSI. *J. Atmos. Oceanic Technol.*, **34**, 797–815, https://doi.org/10.1175/JTECH-D-16-0136.1.

Mason, I. B., 2003: Binary events. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*, I. T. Jolliffe and D. B. Stephenson, Eds., John Wiley and Sons, 37–76.

McNally, A. P., 2009: The direct assimilation of cloud-affected satellite infrared radiances in the ECMWF 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **135**, 1214–1229, https://doi.org/10.1002/qj.426.

——, P. D. Watts, J. A. Smith, R. Engelen, G. A. Kelly, J. N. Thépaut, and M. Matricardi, 2006: The assimilation of AIRS radiance data at ECMWF. *Quart. J. Roy. Meteor. Soc.*, **132**, 935–957, https://doi.org/10.1256/qj.04.171.

Minamide, M., and F. Zhang, 2017: Adaptive observation error inflation for assimilating all-sky satellite radiance. *Mon. Wea. Rev.*, **145**, 1063–1081, https://doi.org/10.1175/MWR-D-16-0257.1.

——, and ——, 2018: Assimilation of all-sky infrared radiances from *Himawari-8* and impacts of moisture and hydrometer initialization on convection-permitting tropical cyclone prediction. *Mon. Wea. Rev.*, **146**, 3241–3258, https://doi.org/10.1175/MWR-D-17-0367.1.

——, and ——, 2019: An adaptive background error inflation method for assimilating all-sky radiances. *Quart. J. Roy. Meteor. Soc.*, **145**, 805–823, https://doi.org/10.1002/qj.3466.

Montmerle, T., F. Rabier, and C. Fischer, 2007: Relative impact of polar-orbiting and geostationary satellite radiances in the Aladin/France numerical weather prediction system. *Quart. J. Roy. Meteor. Soc.*, **133**, 655–671, https://doi.org/10.1002/qj.34.

Nakanishi, M., 2001: Improvement of the Mellor-Yamada turbulence closure model based on large-eddy simulation data. *Bound.-Layer Meteor.*, **99**, 349–378, https://doi.org/10.1023/A:1018915827400.

——, and H. Niino, 2006: An improved Mellor-Yamada level-3 model: Its numerical stability and application to a regional prediction of advection fog. *Bound.-Layer Meteor.*, **119**, 397–407, https://doi.org/10.1007/s10546-005-9030-8.

Okamoto, K., 2013: Assimilation of overcast cloudy infrared radiances of the geostationary MTSAT-1R imager. *Quart. J. Roy. Meteor. Soc.*, **139**, 715–730, https://doi.org/10.1002/qj.1994.

Otkin, J. A., 2010: Clear and cloudy-sky infrared brightness temperature assimilation using an ensemble Kalman filter. *J. Geophys. Res.*, **115**, D19207, https://doi.org/10.1029/2009JD013759.

——, 2012a: Assessing the impact of the covariance localization radius when assimilating infrared brightness temperature observations using an ensemble Kalman filter. *Mon. Wea. Rev.*, **140**, 543–561, https://doi.org/10.1175/MWR-D-11-00084.1.

——, 2012b: Assimilation of water vapor sensitive infrared brightness temperature observations during a high impact weather event. *J. Geophys. Res.*, **117**, D19203, https://doi.org/10.1029/2012JD017568.

——, and R. Potthast, 2019: Assimilation of all-sky SEVIRI infrared brightness temperatures in a regional-scale ensemble data assimilation system. *Mon. Wea. Rev.*, **147**, 4481–4509, https://doi.org/10.1175/MWR-D-19-0133.1.

——, ——, and A. S. Lawless, 2018: Nonlinear bias correction for satellite data assimilation using Taylor series polynomials. *Mon. Wea. Rev.*, **146**, 263–285, https://doi.org/10.1175/MWR-D-17-0171.1.

Pavelin, E. G., S. J. English, and J. R. Eyre, 2008: The assimilation of cloud-affected infrared satellite radiances for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.*, **134**, 737–749, https://doi.org/10.1002/qj.243.

Pleim, J. E., 2007: A combined local and nonlocal closure model for the atmospheric boundary layer. Part I: Model description and testing. *J. Appl. Meteor. Climatol.*, **46**, 1383–1395, https://doi.org/10.1175/JAM2539.1.

Qin, Z. K., X. L. Zou, and F. Z. Weng, 2013: Evaluating added benefits of assimilating GOES imager radiance data in GSI

for coastal QPFs. *Mon. Wea. Rev.*, **141**, 75–92, https://doi.org/10.1175/MWR-D-12-00079.1.

——, ——, and ——, 2017: Impacts of assimilating all or GOES-like AHI infrared channels radiances on QPFs over eastern China. *Tellus*, **69A**, 1345265, https://doi.org/10.1080/16000870.2017.1345265.

Sawada, Y., K. Okamoto, M. Kunii, and T. Miyoshi, 2019: Assimilating every-10-minute Himawari-8 infrared radiances to improve convective predictability. *J. Geophys. Res. Atmos.*, **124**, 2546–2561, https://doi.org/10.1029/2018JD029643.

Schmit, T. J., M. M. Gunshor, W. P. Menzel, J. J. Gurka, J. Li, and A. S. Bachmeier, 2005: Introducing the next-generation advanced baseline imager on GOES-R. *Bull. Amer. Meteor. Soc.*, **86**, 1079–1096, https://doi.org/10.1175/BAMS-86-8-1079.

——, P. Griffith, M. M. Gunshor, J. M. Daniels, S. J. Goodman, and W. J. Lebair, 2017: A closer look at the ABI on the GOES-R series. *Bull. Amer. Meteor. Soc.*, **98**, 681–698, https://doi.org/10.1175/BAMS-D-15-00230.1.

Skamarock, W. C., and Coauthors, 2008: A description of the Advanced Research WRF version 3. NCAR Tech. Note NCAR/TN-475+STR, 113 pp., https://doi.org/10.5065/D68S4MVH.

Snook, N., M. Xue, and Y. Jung, 2011: Analysis of a tornadic mesoscale convective vortex based on ensemble Kalman filter assimilation of CASA X-band and WSR-88D radar data. *Mon. Wea. Rev.*, **139**, 3446–3468, https://doi.org/10.1175/MWR-D-10-05053.1.

Snyder, C., and F. Zhang, 2003: Assimilation of simulated Doppler radar observations with an ensemble Kalman filter. *Mon. Wea. Rev.*, **131**, 1663–1677, https://doi.org/10.1175//2555.1.

Sobash, R. A., and D. J. Stensrud, 2013: The impact of covariance localization for radar data on EnKF analyses of a developing MCS: Observing system simulation experiments. *Mon. Wea. Rev.*, **141**, 3691–3709, https://doi.org/10.1175/MWR-D-12-00203.1.

Stengel, M., P. Undén, M. Lindskog, P. Dahlgren, N. Gustafsson, and R. Bennartz, 2009: Assimilation of SEVIRI infrared radiances with HIRLAM 4D-Var. *Quart. J. Roy. Meteor. Soc.*, **135**, 2100–2109, https://doi.org/10.1002/qj.501.

Stensrud, D. J., and Coauthors, 2013: Progress and challenges with Warn-on-Forecast. *Atmos. Res.*, **123**, 2–16, https://doi.org/10.1016/j.atmosres.2012.04.004.

Sun, J., and Coauthors, 2014: Use of NWP for nowcasting convective precipitation: Recent progress and challenges. *Bull. Amer. Meteor. Soc.*, **95**, 409–426, https://doi.org/10.1175/BAMS-D-11-00263.1.

Szyndel, M. D. E., G. Kelly, and J. N. Thépaut, 2005: Evaluation of potential benefit of assimilation of SEVIRI water vapour radiance data from Meteosat-8 into global numerical weather prediction analyses. *Atmos. Sci. Lett.*, **6**, 105–111, https://doi.org/10.1002/asl.98.

Thompson, G., P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, https://doi.org/10.1175/2008MWR2387.1.

Tong, C.-C., Y. Jung, M. Xue, and C. Liu, 2020: Direct assimilation of radar data with ensemble Kalman filter and hybrid ensemble-variational method in the National Weather Service operational data assimilation system GSI for the stand-alone regional FV3 model at a convection-allowing resolution. *Geophys. Res. Lett.*, **47**, e2020GL090179, https://doi.org/10.1029/2020GL090179.

Tong, M., and M. Xue, 2005: Ensemble Kalman filter assimilation of Doppler radar data with a compressible nonhydrostatic model: OSS experiments. *Mon. Wea. Rev.*, **133**, 1789–1807, https://doi.org/10.1175/MWR2898.1.

Wang, C., and X. Huang, 2014: Parallax correction in the analysis of multiple satellite data sets. *IEEE Geosci. Remote Sens. Lett.*, **11**, 965–969, https://doi.org/10.1109/LGRS.2013.2283573.

Wang, Y., Z. Liu, S. Yang, J. Min, L. Chen, Y. Chen, and T. Zhang, 2018: Added value of assimilating Himawari-8 AHI water vapor radiances on analyses and forecasts for "7.19" severe storm over North China. *J. Geophys. Res. Atmos.*, **123**, 3374–3394, https://doi.org/10.1002/2017JD027697.

Wheatley, D. M., K. H. Knopfmeier, T. A. Jones, and G. J. Creager, 2015: Storm-scale data assimilation and ensemble forecasting with the NSSL experimental Warn-on-Forecast System. Part I: Radar data experiments. *Wea. Forecasting*, **30**, 1795–1817, https://doi.org/10.1175/WAF-D-15-0043.1.

Whitaker, J. S., and T. M. Hamill, 2002: Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, **130**, 1913–1924, https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2.

——, and ——, 2012: Evaluating methods to account for system errors in ensemble data assimilation. *Mon. Wea. Rev.*, **140**, 3078–3089, https://doi.org/10.1175/MWR-D-11-00276.1.

Yang, C., Z. Liu, F. Gao, P. P. Childs, and J. Min, 2017: Impact of assimilating GOES imager clear-sky radiance with a Rapid Refresh assimilation system for convection-permitting forecast over Mexico. *J. Geophys. Res. Atmos.*, **122**, 5472–5490, https://doi.org/10.1002/2016JD026436.

Yang, J., Z. Zhang, C. Wei, F. Lu, and Q. Guo, 2017: Introducing the new generation of Chinese geostationary weather satellites, FengYun-4. *Bull. Amer. Meteor. Soc.*, **98**, 1637–1658, https://doi.org/10.1175/BAMS-D-16-0065.1.

Yussouf, N., D. C. Dowell, L. J. Wicker, K. H. Knopfmeier, and D. M. Wheatley, 2015: Storm-scale data assimilation and ensemble forecasts for the 27 April 2011 severe weather outbreak in Alabama. *Mon. Wea. Rev.*, **143**, 3044–3066, https://doi.org/10.1175/MWR-D-14-00268.1.

Zhang, F., C. Snyder, and J. Sun, 2004: Impacts of initial estimate and observations on the convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, **132**, 1238–1253, https://doi.org/10.1175/1520-0493(2004)132<1238:IOIEAO>2.0.CO;2.

——, M. Minamide, and E. E. Clothiaux, 2016: Potential impacts of assimilating all-sky infrared satellite radiances from GOES-R on convection-permitting analysis and prediction of tropical cyclones. *Geophys. Res. Lett.*, **43**, 2954–2963, https://doi.org/10.1002/2016GL068468.

——, ——, R. G. Nystrom, X. Chen, S.-J. Lin, and L. M. Harris, 2019: Improving Harvey forecasts with next-generation weather satellites: Advanced hurricane analysis and prediction with assimilation of GOES-R all-sky radiances. *Bull. Amer. Meteor. Soc.*, **100**, 1217–1222, https://doi.org/10.1175/BAMS-D-18-0149.1.

Zhang, Y., F. Zhang, and D. J. Stensrud, 2018: Assimilating all-sky infrared radiances from *GOES-16* ABI using an ensemble Kalman filter for convection-allowing severe thunderstorms prediction. *Mon. Wea. Rev.*, **146**, 3363–3381, https://doi.org/10.1175/MWR-D-18-0062.1.

——, D. J. Stensrud, and F. Zhang, 2019: Simultaneous assimilation of radar and all-sky satellite infrared radiance observations for convection-allowing ensemble analysis and prediction

of severe thunderstorms. *Mon. Wea. Rev.*, **147**, 4389–4409, https://doi.org/10.1175/MWR-D-19-0163.1.

——, ——, and E. E. Clothiaux, 2021: Benefits of the Advanced Baseline Imager (ABI) for ensemble-based analysis and prediction of severe thunderstorms. *Mon. Wea. Rev.*, **149**, 313–332, https://doi.org/10.1175/MWR-D-20-0254.1.

Zhu, K., M. Xue, Y. Pan, M. H. S. G. Benjamin, S. S. Weygandt, and H. Lin, 2019: The impact of satellite radiance data assimilation within a frequently updated regional forecast system using GSI-based ensemble Kalman filter. *Adv. Atmos. Sci.*, **36**, 1308–1326, https://doi.org/10.1007/s00376-019-9011-3.

Zhu, Y., and Coauthors, 2016: All-sky microwave radiance assimilation in NCEP's GSI analysis system. *Mon. Wea. Rev.*, **144**, 4709–4735, https://doi.org/10.1175/MWR-D-15-0445.1.

Zou, X., Z. Qin, and F. Weng, 2011: Improved coastal precipitation forecasts with direct assimilation of *GOES-11/12* imager radiances. *Mon. Wea. Rev.*, **139**, 3711–3729, https://doi.org/10.1175/MWR-D-10-05040.1.

——, ——, and Y. Zheng, 2015: Improved tropical storm forecasts with *GOES-13/15* imager radiance assimilation and asymmetric vortex initialization in HWRF. *Mon. Wea. Rev.*, **143**, 2485–2505, https://doi.org/10.1175/MWR-D-14-00223.1.